DECONSTRUCTING THE RIBOSOME:

SPECIFIC INTERACTIONS OF A RIBOSOMAL RNA FRAGMENT WITH INTACT AND FRAGMENTED L23 RIBOSOMAL PROTEIN

A Thesis
Presented to
The Academic Faculty

by

Poorna Roy

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
May 2013

DECONSTRUCTING THE RIBOSOME:

SPECIFIC INTERACTIONS OF A RIBOSOMAL RNA FRAGMENT

WITH INTACT AND FRAGMENTED L23 RIBOSOMAL PROTEIN

Approved by:

Dr. Loren D Williams, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Nicholas V Hud
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Adegboyega K Oyelere
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Roger M Wartell
School of Biology
Georgia Institute of Technology

Date Approved: March 25, 2013

This thesis is dedicated to my parents

Late Mr.Dipak Ranjan Ray & Late Mrs. Chandana Roy

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my deep sense of gratitude and respect to my supervisor Dr. Loren Williams, for his constant guidance and encouragement throughout my project period and for providing the necessary facilities. I would also like to thank the members of thesis committee, Dr. Nicholas V Hud, Dr. Adegboyega K Oyelere and Dr. Roger M Wartell for their constructive feedback on my research.

I am grateful to all members in the Williams lab for making me feel a part of the team and for their unfailing support and encouragement. I am particularly grateful to Dana Schneider and Jessica Bowman for mentoring me and furthering my technical capabilities.

Finally, I thank my family, my dada, bourani, my uncle and my friends Debarati, Vishal, Tanushree, Shauvik and PrithaDi for believing in me and helping get through some very difficult times.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviation or Symbol	Term
Å	Angstrom
μg	Microgram
μL	Microliter
μM	Micromolar
3-AT	3-Amino-1,2,4-triazole
g	Gram
MDa	Mega Daltons
mg	Miligram
mM	Millimolar
nm	Nanometer
SHAPE	Selective 2′-Hydroxyl Acylation analyzed by Primer Extension

LIST OF NOMENCLATURE

Abbreviation	Term
DIII	Domain III
$DIII^{core}$	Domain III-core
GAD	Gal4 activation domain
H-bond	Hydrogen bond
HIS3	Gene encoding enzyme catalyzing sixth step in histidine biosynthesis
$L23^{peptide}$	Peptide derived from ribosomal protein L23
LSU	Ribosome Large subunit
LUCA	Last Universal Common Ancestor
Mg^{2+}-μC	Magnesium microcluster
mRNA	Messenger RNA
PTC	Peptidyl Transferase Center
rRNA	Ribosomal RNA
rProtein	Ribosomal protein
SSU	Ribosome Small subunit
tRNA	Transfer RNA
Y3H	Yeast three-hybrid assay

SUMMARY

The complexity of translation is a classical dilemma in the evolution of biological systems. Efficient translation requires coordination of complex, highly evolved RNAs and proteins; however, complex, highly evolved RNAs and proteins could not evolve without efficient translation system. At the heart of this complexity is the ribosome, itself a remarkably complex molecular machine. Our work illustrates the ribosome as deconstructed units of modification.

Here we have deconstructed a segment of the ribosome to interacting RNA-protein units. L23 interacts *in vivo* with both Domain III (DIII) and Domain III^{core} (DIII^{core}) independently of the fully assembled ribosome. This suggests that DIII^{core} represents the functional rRNA unit in DIII-L23 interaction. Furthermore, L23^{peptide} sustains binding function *in vitro* with both DIII and DIII^{core} independently of any stabilizing effects from the globular domain of L23. The ability of L23^{peptide} to form a 1:1 complex with both DIII and DIIIcore suggests that L23peptide is the functional rProtein unit in DIII-L23 interaction. We believe that our results will stimulate interest and discussions in the significance of 3D architecture and units of evolution in the ribosome. The ubiquity of the ribosome in cellular life prognosticates that our results impact and appeal to biologists, chemists, bioinformaticists, as well as the general scientific community.

CHAPTER 1

INTRODUCTION

Evolution of translation

The origin of complexity in biological systems is a classic evolutionary dilemma. Storage and transfer of genetic information and efficient functioning of metabolic machinery are two necessary conditions for the origin and sustenance of life. Translation bridges the two necessary conditions by employing nucleic acids that encode genetic information and result in the synthesis of proteins. On a molecular level, translation is perhaps the most complex of biological processes. The universal macromolecule that serves as the site for translation is the ribosome. Messenger RNA (mRNA), transcribed from the organism's genome, binds with the ribosomal small and large subunits to initiate protein biosynthesis (Figure 1.1).

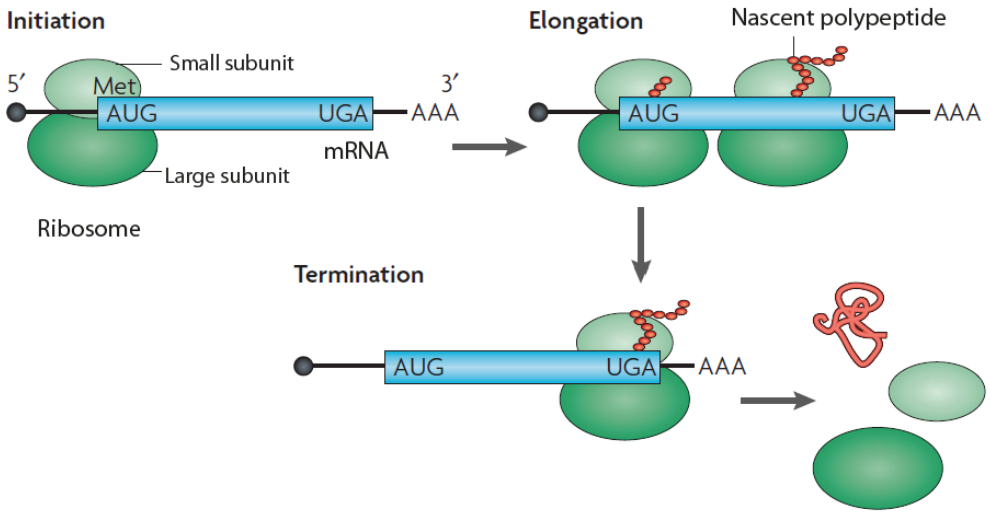


Figure 1.1. Schematic representation of translation of mRNA by ribosome. [Adapted from [1]]

The ribosome is a ribonucleoprotein complex. Ribosomal RNA (rRNA) catalyzes peptide bond formation while ribosomal proteins (rProteins) are involved in stabilizing the overall structure of the ribosome as well as in the basic steps of translation such as

initiation and elongation [2]. Early ideas about the origin of translation baffled scientists with the classic chicken and egg problem: If the early ribosome required proteins for its own function, how could it evolve in an "RNA world"? [3]. The discovery of ribozymes, catalytic RNAs [4] provided evidence that the early ribosomes could have been composed solely of RNA [5]. According to the "RNA world" model, life on Earth evolved on the basis of RNA replication, which was followed by protein synthesis [5]. Before the emergence of proteins, RNA would have had to carry out both coding and catalysis [6,7]. Proposed theories suggest the ribosome emerged as a simple primitive ribozyme that catalyzed non-coded peptide synthesis [8,9] and then enhanced in rapidity and accuracy with time and increasing complexity [10].

The ribosome is an intricately complex structure. Such complexity raises questions about the emergence of life in a pre-biotic world. If translation involved RNA as the sole catalyst, molecules of modest structural complexity would have populated the ancient world [11]. Therefore, the ribosome and its associated tRNAs and mRNAs could have evolved in a step-wise fashion from small, preexisting RNAs, which had other ancient functions [3].

Alternative evolutionary theories suggest a peptide-dominated ancient world [12,13]. Recent supporting evidence for a peptide-dominated ancient world includes ancient lineage of some proteins and a lack of RNA interactions for most proteins [14,15]. In translation, RNA and protein are fully interdependent in a process of molecular complexity. Origins and evolution of RNA/protein interdependence in the ribosome followed a sequence of incremental and elementary steps, each linked to immediate fitness, without foresight.

Ribosome as a molecular fossil

The ribosome is responsible for protein synthesis in all living organisms and is our most direct macromolecular connection to distant evolutionary past and early life [16-18]. Our current understanding about origin of life suggests that biological materials are derived from a bottom-up process, with gradual chemical evolution leading to the emergence of molecular networks of increasingly greater complexity [19].

Several complex organic molecules failed to achieve self-replication, thereby being led to extinction (Figure 1.2). Among macromolecules, amino acid and nucleic acid polymers determined the critical threshold for the emergence of life. An RNA world model proposed self-replicating RNA subsequently combining with proteins in an RNA-protein (RNP) world. Transition from the RNP world to DNA based life was relatively recent; the last universal common ancestor (LUCA) is posited to have had a DNA genome, using RNA and proteins for biocatalysis [20].

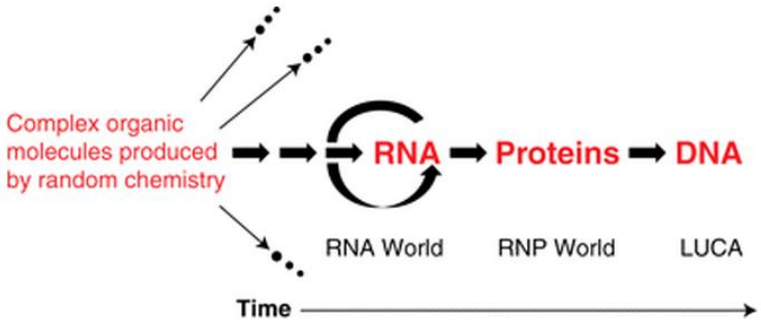


Figure 1.2. An RNA world model for the successive appearance of RNA, proteins, and DNA during the evolution of life on Earth. Figure reproduced from [20].

RNA and protein cooperate in the modern ribosome [21]. A modern RNA, such as rRNA, is a potential relic or a fossil of an ancient RNA if it fulfills one or more of the following criteria: (i) is catalytic, (ii) is ubiquitous, (iii) has a central role in metabolism, (iv) carries out a role served by a protein in other organisms [7].

Identifying such “fossilized” rRNA in eukaryotes allows reconstruction of the ancient ribosome that facilitated translation in the LUCA [22]. LUCA possessed

universal genes involved in translation, transcription, RNA processing and degradation, intermediary and RNA metabolism, and compartmentalization [23-25].

The availability of high-resolution three dimensional structures of ribosomes from distant regions of the evolutionary tree has allowed sequence and conformational comparisons between rRNAs from different domains of life. Much of the diversity of conformation and sequence between bacterial and archaeal ribosomes predates the LUCA. Recent studies on the 3D structures of ribosome suggest that a near-modern ribosome was well established before the LUCA, beyond the root of the phylogenic tree [26].

Comparison of rRNA sequences from different domains showed phylogenetic conservation in the peptidyl transferase center (PTC) where catalysis occurs. The PTC lacks protein; catalytic activity of the ribosome is mediated by the rRNA. The highly conserved PTC and the dense RNA active site at the core of the ribosome argue strongly in favor of the ribosome being a molecular fossil [27].

A structure-based and sequence-based comparison of the large subunits (LSUs) of *Haloarcula marismortui* and *Thermus thermophilus* shows that sequence and conformational similarity of the 23S rRNAs are greatest near the PTC core, then diverge with increasing distance from the PTC (Figure 1.3) [18]. The results suggest that the conformation and interactions of both RNA and protein can be described as changing, in an observable manner, over evolutionary time, thus making the ribosome an ancient molecular fossil [18].

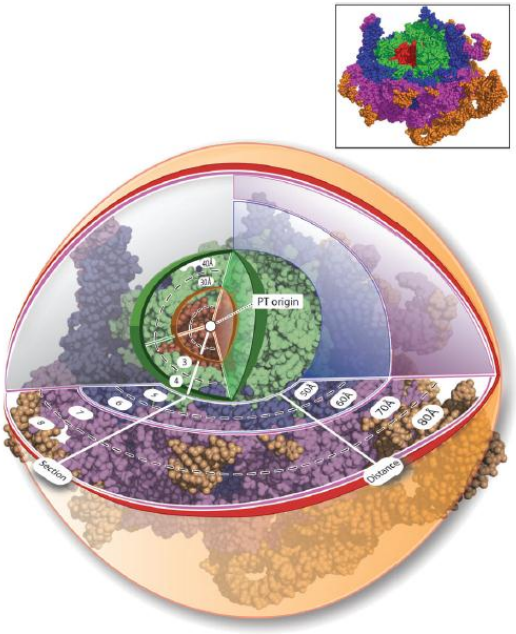


Figure 1.3. Peeling the ribosomal onion. The *Haloarcula marismortuii* (and *Thermus thermophilus*, not shown) LSUs have been sectioned into concentric shells, with the origin at the site of peptidyl transfer (the PT-origin) [18].

All parts of the ribosome are not of the same age

The complex macromolecular structure of the ribosome suggests that sequential addition of structural components over time led to the emergence of the modern ribosome. rRNA molecules associate with rProteins to form ribosomal subunits and such an assembly has been verified by numerous reconstitution experiments [28-34].

During reconstitution experiments, a specific order was observed in which the different ribosomal components interact with each other. The assembly order of ribosomal components may recapitulate ribosomal evolution because components required during initial assembly are more likely to have evolved first. The *in vitro* assembly map illustrates known dependencies of ribosomal components on one another (Figure 1.4) [35]. On the basis of the 30S and 50S assembly maps, ribosome assembly is believed to be cooperative, where binding of ribosomal proteins induces structural changes in the rRNA. Some ribosomal proteins bind directly to the RNA whereas others

are only incorporated later after other proteins have been added. *In vivo* assembly of the ribosomal components is likely to follow the same order [36].

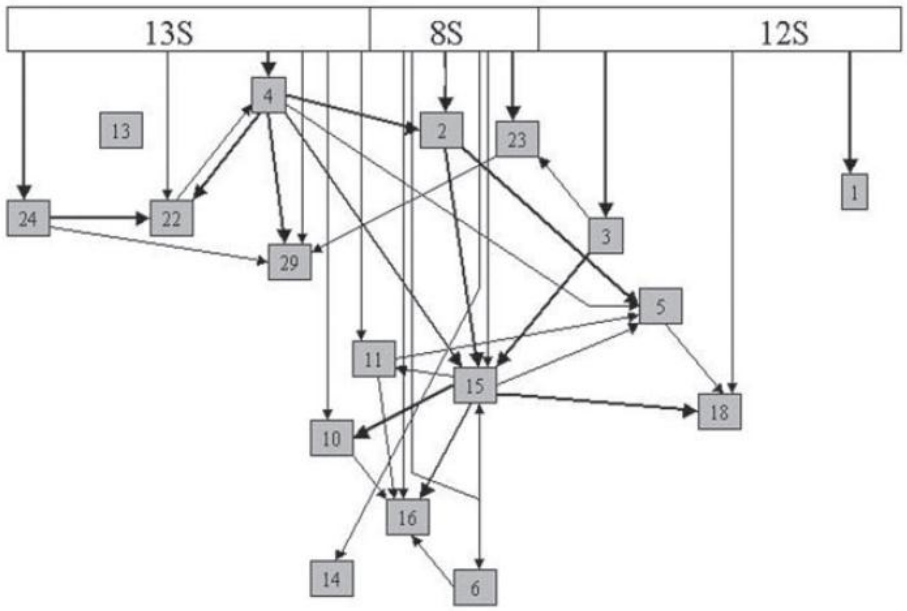


Figure 1.4. Location of universal proteins in the 50S assembly map. The 23S rRNA is represented as three fragments of 13S, 8S and 12S. The proteins that have equivalents in all three Domains of life are shown in this map exclusively as compared to the original assembly map[35]. [Adapted from [36]]

Not all the proteins that are required for the assembly of the modern ribosome are of the same age. Though many rProteins are universally conserved and can be traced back to the LUCA, some of the proteins are exclusive to only one domain of life and therefore believed to be of relatively modern origin. To determine the hierarchy of protein binding to rRNA, Nomura and co-workers (for the SSU) [37] and Nierhaus and co-workers (for the LSU) [35] varied the order in which rProteins were added during *in vitro* assembly. The thermodynamical interdependence of protein-binding was related to the chronological ordering of the proteins. A protein is considered terminal if no other protein depends on its presence for assembly. Based on the assembly map and the order observed during reconstitution of the ribosome, earliest proteins in ribosomal reconstitution were found to be earlier and more important additions to the ribosomal assembly [36].

Structure of the modern ribosome provides evidence for RNA replication predating protein synthesis. It is likely that rRNA evolved from small, self-replicating RNAs to large complex rRNA in the modern ribosome by conflation of originally independent functional fragments [38]. Subsequent studies support that different parts of the rRNA are of different ages and propose a stepwise addition of structural elements [17,18,39-41] and varied evolutionary rates of such structural elements [42].

Based on molecular interactions [17,41], distance from the PTC [18] and rRNA sequence conservation in the three domains of life [43], different models proposed for the evolution of 23S rRNA (Figure 1.5) postulate chronology of molecular events in early life. The models postulate that the oldest parts of rRNA would be integral to the evolving translational machinery and would exhibit high interconnectivity [44] and sequence conservation in all domains of life [43]

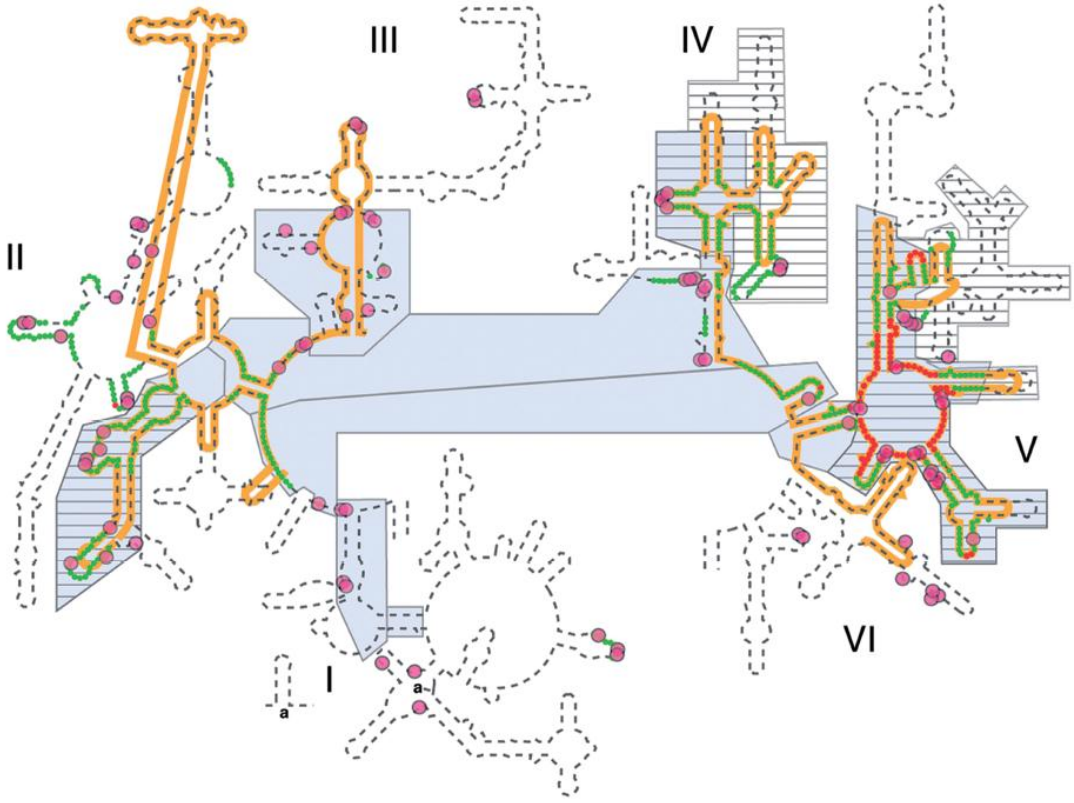


Figure 1.5. Various models of evolution of 23S rRNA. The dashed line illustrates the canonical secondary structure of the *T. thermophilus* 23S rRNA. Secondary structural domains are indicated by roman numerals. The red and green lines show the two inner

shells of the ribosomal onion of Hsiao and Williams, marking the rRNA that is in closest proximity, in three dimensions, to the site of peptidyl transfer. The gray boxes are ancient according to Steinberg's A-minor model. The hashed boxes (with black horizontal lines) are ancient according to the networking analysis of Fox. Multidentate Mg^{2+}-phosphate interactions, also proposed as an indicator of ancient rRNA, are indicated by magenta circles. The orange line shows the universally conserved portions of the 23S rRNA in bacteria, archaea, eukarya, and in mitochondria, as determined by Gutell and Harvey [45].

Among the ribosomal components that are highly conserved throughout extant life [40], PTC is considered to be one of the oldest structures in biology, predating coded protein [10,39,46-48]. The striking similarity of PTC in all living organisms has led to the inference that the PTC must have been formed prior to the division of tree of life into three major domains [49,50]. Since the PTC is contained in Domain V of 23S rRNA, Domain V is expected to be one of the oldest parts of the ribosome [51] and shows extensive interconnectivity with other secondary domains of 23S rRNA. Comparative structural methods and analysis of conserved sequences of rRNA have led to different models of ancient ribosomal evolution, all of which agree on LSU being the oldest in evolution near the PTC and younger near the surface [18,26].

Domain II and Domain IV show high interconnectivity with each other and with Domain V. Consequently, they are believed to be amongst the early components of the rRNA. Though Domain II and Domain IV have comparable interconnectivity with other secondary domains of 23S rRNA, majority of the molecular interactions between 16S rRNA and 23S rRNA are through Domain IV. The lack of proximity of Domain II with tRNAs and minimal association with the 30S subunit suggests that Domain II might be a relatively recent to Domain IV [44].

CHAPTER 2

ARCHITECTURE AND FUNCTIONS OF RIBOSOME

The ribosome is responsible for protein synthesis in all organisms. The structure of the ribosome does vary significantly with the size of prokaryotic and eukaryotic ranging from molecular weights of 2.5 to 4 MDa (for prokaryotes and eukaryotes, respectively) [52]. Ribosomes typically contain 50 to 60 percent RNA as an integral part of their structures. The large subunit (LSU) contains the 5S rRNA and the 23S rRNA which houses the peptidyl transfer center that catalyzes peptide bond formation. The small subunit (SSU) contains the 16S rRNA which mediates recognition between mRNA codons and tRNA anticodons [53]. The 16S rRNA of the SSU is associated with 20 or more rProteins depending on the species [54]; the LSU is associated with 50 or more rProteins (Figure 2.1).

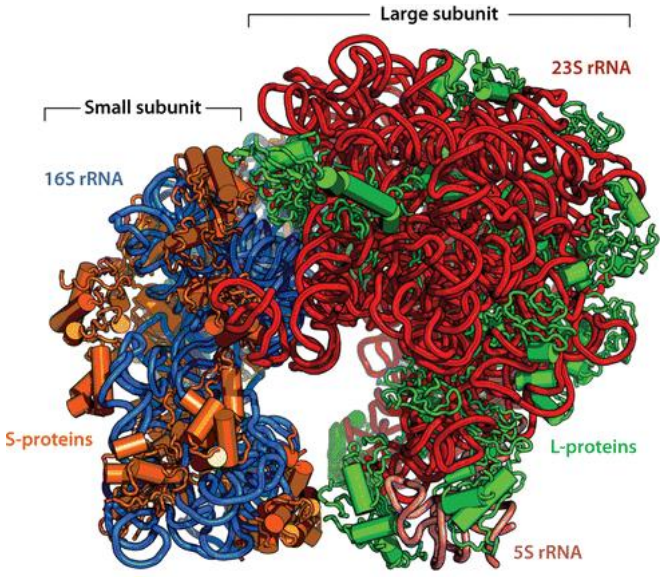


Figure 2.1. The *E. coli* 70S ribosome. 16S rRNA is depicted in blue and the 23S rRNA in red. The rProteins associated with SSU and LSU are shown in orange and green respectively [55].

The SSU and LSU associate by intersubunit bridges during translation to form the active ribosome. In order to translate the genetic information stored in the mRNA, the ribosome selects cognate tRNA that can base pair with a specific mRNA codon. This process is called “decoding” [56]. A V-shaped cavity in the middle of the large subunit that harbors the PTC, also houses the highly conserved A- and P-loops. The A- and P-loops accommodate the 3’-termini (CCA) of the A (aminoacyl) and the P (peptidyl) tRNAs. During protein biosynthesis the A-site tRNA attached to the nascent polypeptide chain, passes into the P-site. Once a peptide bond is formed, the deacylated P-site tRNA acts as the “leaving group” and moves from the P-site to the E (exit)-site [57].

Factors stabilizing ribosomal architecture

The ribosome is composed of simpler structural motifs that are held together by stabilizing interactions that lead to the intricate architecture of the ribosome. Structure, like sequence, reveals information about macromolecular origins and evolution. During RNA evolution, the structure is much more conserved than the sequence [58]. A variety of nucleotide sequence of rRNA can give rise to homologous secondary and three dimensional structures which have the same molecular function [49]. In the analysis of ribosomal architecture, identification of small structural motifs and the factors that contribute to their organization into larger subassemblies is of fundamental interest.

There are three major factors that aid the assembly of structural motifs in the ribosome: (i) metal ions that mediate short and long-range tertiary interactions (ii) RNA-RNA interaction such as base-pairing of complementary nucleotides and base-phosphate interaction (iii) RNA-protein interaction

Metal ions in the ribosome

RNAs associate with metal ions [59]. When large RNAs fold into compact structures, negatively charged phosphate groups are brought into close proximity. Cations help neutralize the negative charge density associated with the RNA phosphate backbone. Such charge neutralization occurs with cationic polyamines as well as inorganic cations [60]. Divalent cations also aid in RNA self-cleavage in ribozymes [61,62]. Mg^{2+} is particularly suitable divalent cation to facilitate rRNA compaction and catalysis because: (a) it is the most abundant intracellular multivalent cation and (b) it has the highest charge density of all biologically available ions, owing to its relatively small ionic radius (0.6 Å) [63]. Mg^{2+} associates preferentially with oxyanions of RNA phosphates over base and ribose atoms, pre-dominantly forming mono and bi-dentate complexes.

Seventy-one magnesium ions associate with the P4–P6 domain of the *Tetrahymena* group 1 intron (PDB 1HR2) [64,65] of which twenty-six contain phosphate oxyanions within their first coordination shell. When two oxyanions enter the first coordination shell of Mg^{2+}, the bidentate chelation of magnesium by RNA phosphates leads to a structure called a bidentate clamp (Figure 2.2A) [66]. A 10-membered ring characterizes these bidentate RNA clamps. Such Mg^{2+} mediated bidentate RNA clamps occur frequently in large RNAs. In the 23S rRNA of *H. marismortui* (PDB entry 1JJ2), 98 out of 118 magnesium ions associate with , phosphate oxyanions within their first coordination shells (characterized by Mg^{2+}–OP distances < 2.4 A˚), frequently forming Mg^{2+} bidentate clamps [63,67] (Figure 2.2B).

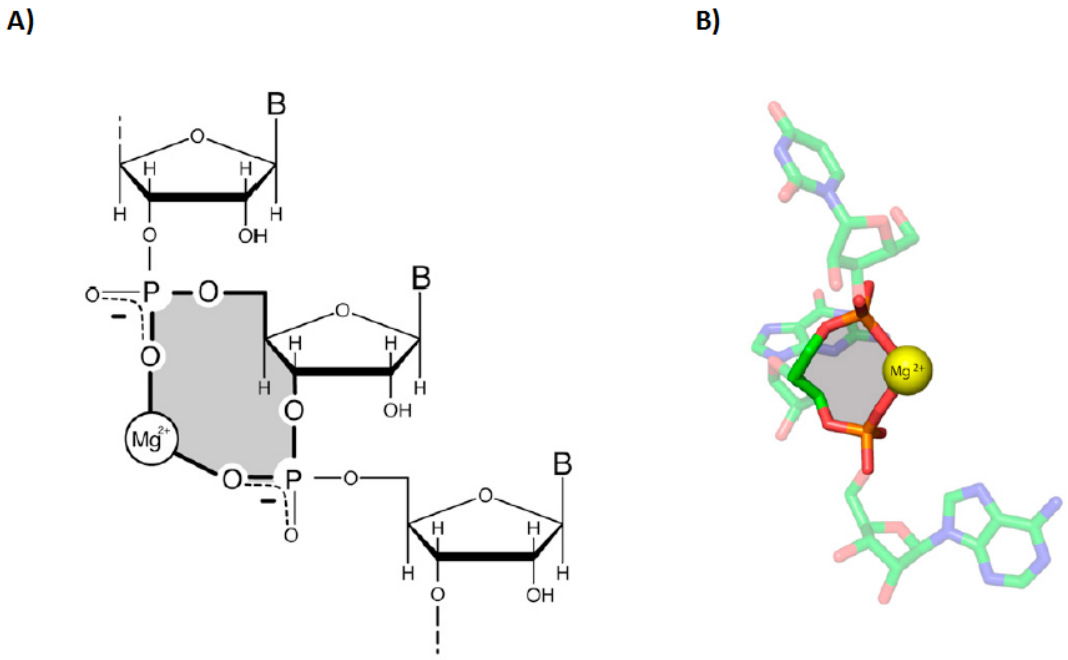


Figure 2.2. Mg^{2+} stabilized interactions in the ribosome. A) Schematic diagram of a bidentate RNA clamp of magnesium, formed when adjacent phosphate groups enter the first coordination shell of a common magnesium ion B) A bidentate RNA clamp of magnesium observed in the H. marismortui ribosomal LSU crystal structure (Mg 8003 from PDB entry 1JJ2) [66]

RNA-RNA interaction

Key events in translation involve RNA-RNA interaction be it start-site selection by the mRNA, decoding or peptide bond formation in the PTC [68-70]. In order to perform such crucial activities [67,71,72], rRNA need to fold into compact architectures.

RNA molecules need to overcome electrostatic self-repulsion to enable such compact folding. The negative charge of each nucleotide is concentrated on the anionic oxygen atoms of phosphate groups enabling them to form very strong hydrogen bonds (H-bonds) with appropriate donors. Each of the RNA bases (A, C, G and U) comprises multiple H-bond donors that can interact with phosphate groups. Such stabilizing interactions between nucleotide bases and the phosphate backbone moieties are referred

to as ‘base-phosphate’ interactions. They help reduce intra-molecular RNA self-repulsion and stabilize compactly folded structural motifs [73].

The extent and type of hydrogen bonding and base-stacking interactions can determine the architecture of RNA motifs. Both Watson-Crick base-pairing as well as GU Wobble base pair can form helical stems of RNA. Helical segments are interspersed and capped by regions called loops [74]. Single stranded regions of RNA can fold back on itself via regions of complementary base pairs forming a hairpin (Figure 2.3A). Single-stranded regions can also form pseudoknots when base pairs intertwine (Figure 2.3B) [75].

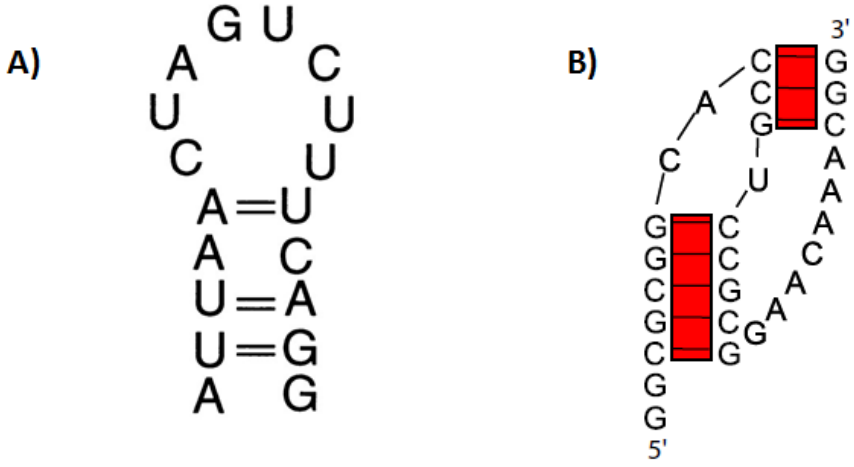


Figure 2.3. RNA structural motifs. A) RNA hairpin, a secondary structural motif B) Loop regions can form long-range interactions as in pseudoknots [Adapted from [75,76]]

Structural motifs of rRNA are arranged in three-dimensional (3D) space by tertiary interactions between the nucleotides of hairpin, internal and junction ‘loops’ of the secondary structure. By establishing local and specific contacts, the tertiary interactions build up 3D structural modules that are characterized by sets of non-Watson-Crick base pairs organized in a precise order [77].

RNA-protein interaction

While rRNAs perform the crucial activities [67,71,72], rProteins charge tRNAs [78], stabilize inter- and intra-subunit ribosomal interactions and assist in ribosomal assembly [35,63]. rProteins counterbalance part of the electrostatic repulsion among RNA nucleotides arising due to the proximity of negative charges .

rProteins interact specifically with rRNA in at least four ways: (i) interactions between the proteins and the edges of bases exposed in the minor grooves of RNA helices, (ii) interactions of proteins with bases that become accessible in widened major grooves of RNA helices (Figure 2.4A), (iii) protein recognition of the flipped out bases of bulged nucleotides, and (iv) insertion of amino acid residues into hydrophobic crevices between exposed nucleotide bases [63]. Additionally, rProteins also interact with the backbone of RNA. For nucleobase-specific interactions, Guanine nucleobase and unpaired RNA structural states are significantly preferred; however nonspecific interactions disfavor guanine, while still favoring unpaired RNA structural states [79].

A recent comparative analysis with the overall RNA surface and RNA-protein interfaces shows that contact surfaces involving RNA motifs have distinctive features that may be useful for the recognition and prediction of interactions [80]. rProteins prefer interaction with structural RNA motifs to that of the complete ribosomal surface. Among RNA motifs, tetraloops show the highest percentage of interaction with rProteins (Figure 2.4B) due to the high shape complementarity of extended tetraloop surface to contacting rProtein.

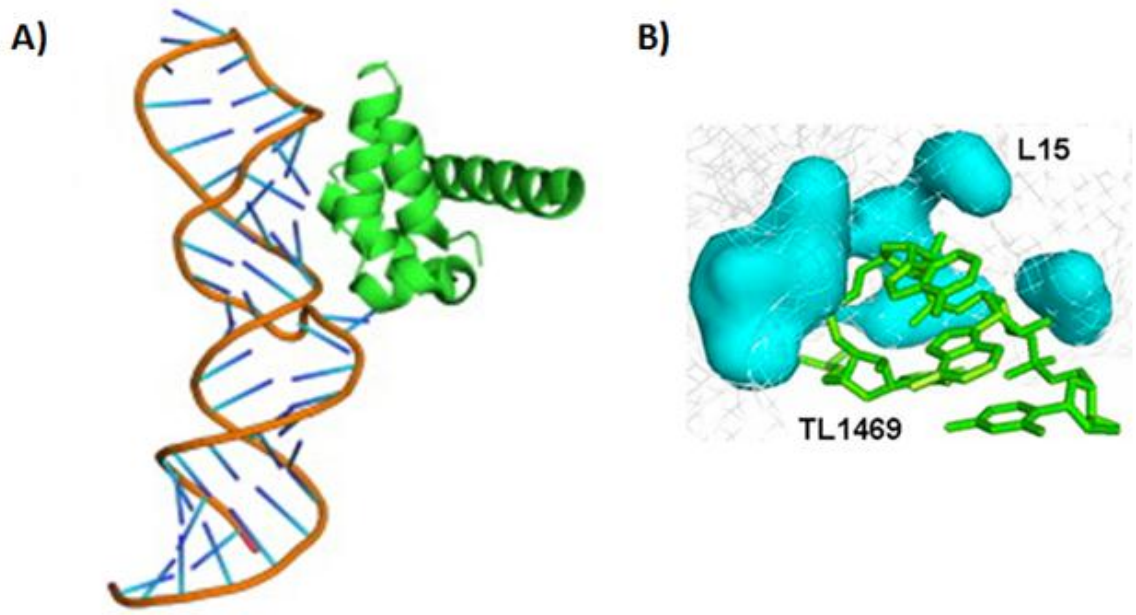


Figure 2.4. Representative RNA-protein complexes. A) RNA-protein complex between Domain IV of 4.5S rRNA and SRP protein [Adapted from [79]]. (B) Tetraloop (TL1469) interaction with rProtein L15 in 50S subunit of T. thermophilus ribosome shows high shape complementarity. [Adapted from [80]]

Architecture of ribosomal RNA

The atomic structures of LSU from *H. morismortui* [67] and *T. thermophilus* [81] revealed six 2° domains in the 23S rRNA (Figure 2.5). In the SSU four 2° domains were found in the 16S rRNA of which, three are considered major domains while one is a minor domain[82]. Each 2° domain of the 16S rRNA folds and assembles with the appropriate ribosomal proteins into a 3D domain, independent of other 2° domains[83-85]. One 3D domain is called the head and others are called the body and the platform [86,87]. The head, body, and platform domains of the SSU have direct functional significance, moving independently during translation[88].

Isolated fragments of 16S rRNA, corresponding closely to the three major domains, interact with ribosomal proteins in specific sites[33,89-92]. The ability of certain rProteins to remain bound to these fragments, or even to rebind to RNA fragments from which proteins have been removed, supports the suggestion that these domains are true structural entities[93].

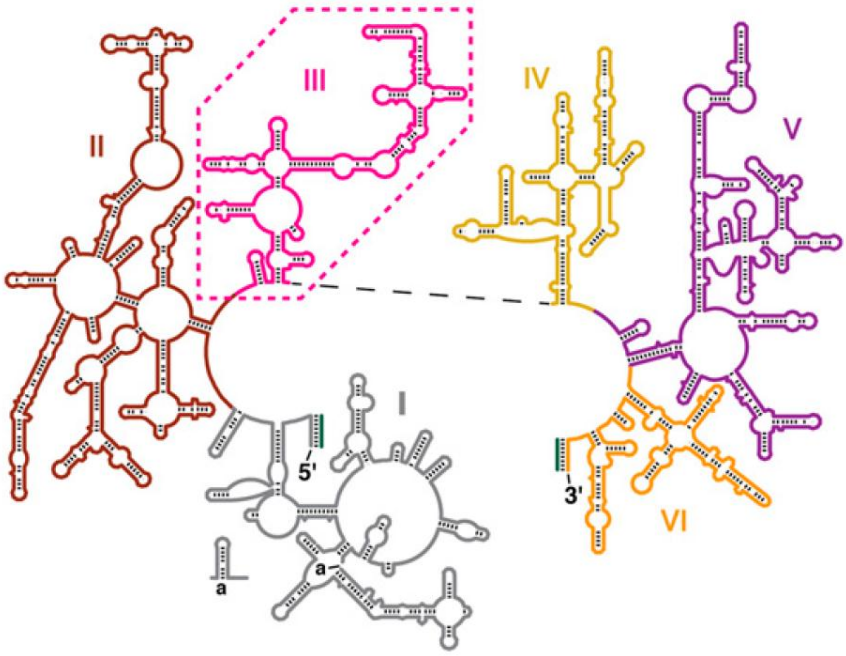


Figure 2.5. Secondary structure of the 23S rRNA of the large subunit of *T. thermophilus* (adapted with permission from Harry Noller). The six secondary structural domains of 23S rRNA are shown: Domain I in gray, Domain II in brown, Domain III in pink, Domain IV in yellow, Domain V in purple, and Domain VI in orange[94].

The 2° domains of 16S rRNA can be segregated into independent and autonomous three-dimensional domains (3D domains) in the assembled SSU. For the 23S rRNA, the general observation has been that it has a very compact structure [49,93,95]. Secondary structure of the LSU is further constrained tertiary interactions [49,96]. The compact structure of the 23S rRNA was confirmed by cryo-EM [53] and X-ray crystallography [67,97] data. The lack of distinct demarcations led to the monolithic appearance of its LSU suggesting the absence of an inherent domain substructure [67].

The colossal structure of the 23S rRNA makes it challenging to comprehend looking at it as a whole. Questions naturally arise as to whether the architectures and early evolution of the SSU and the LSU are fundamentally different, and if so, why? Do isolated 2° domains of the 16S rRNA but not the 23S rRNA act as 3D domains and fold to near-native 3D structures? How are the 2° domains of the 16S and 23S rRNAs related to 3D structure, function, and evolution of the ribosome? The answer is simple: "To convey a sense of how it is put together, the subunit must be dissected into its chemical components."[67]

CHAPTER 3

DECONSTRUCTION OF THE RIBOSOME

Domain III is an autonomously folding domain of 23S rRNA

A recent development in unraveling the architecture of 23S rRNA was made by Athavale *et al*[94]. In an attempt to experimentally probe the domain structure of the LSU, it was shown that one isolated 2° domain of the 23S rRNA, Domain III, can fold to a near-native state in absence of the remainder of the LSU, and appears to be a true 3D domain.

The structure of Domain III rRNA alone (Domain III^{alone}) and when contained within the intact 23S rRNA (Domain III^{23S}) was probed using SHAPE (selective 2'-hydroxyl acylation analyzed by primer extension), in the absence and presence of magnesium. SHAPE exploits the reactivity of the 2'-hydroxyl groups of RNA with electrophilic chemical probing reagents such as NMIA (N-methylisatoic anhydride) or BzCN (benzoyl cyanide)[98,99]. The relative reactivities of the 2'-hydroxyl groups of various nucleotides are sensitive primarily to local RNA flexibility. Consequently, paired nucleotides within helical regions are generally less flexible and less reactive toward SHAPE reagents than unpaired nucleotides. The canonical secondary structure of the 23S rRNA based on comparative sequence analysis [97,100], is strongly supported by previous SHAPE experiments [101]. Results of SHAPE analysis of Domain III^{alone} and Domain III^{23S} in absence and presence of magnesium ions support the hypothesis that Domain III^{alone} folds to a near-native state with secondary structure, intra-domain tertiary interactions, and inter-domain interactions that are independent of whether or not it is embedded in the intact 23S rRNA or within the LSU.

The close correspondence of the SHAPE data of Domain III^{alone} to the canonical secondary structure of Domain III is evident in Figure 3.1, where SHAPE reactivity of Domain III^{alone} is mapped onto the canonical secondary structure. These data were

obtained in presence of 250 mM Na^+ ions and in the absence of divalent cations. Under these conditions, RNA is expected to assume secondary structure but not necessarily tertiary structure[102,103]. Consistent with this tendency, the correspondence between SHAPE reactivities and the secondary structure is very nearly perfect. Nucleotides of Domain III^{alone} were ranked using their absolute SHAPE reactivities relative to A1572 (highest reactivity) and binned into four groups, which are indicated in Figure 3.1 [94].

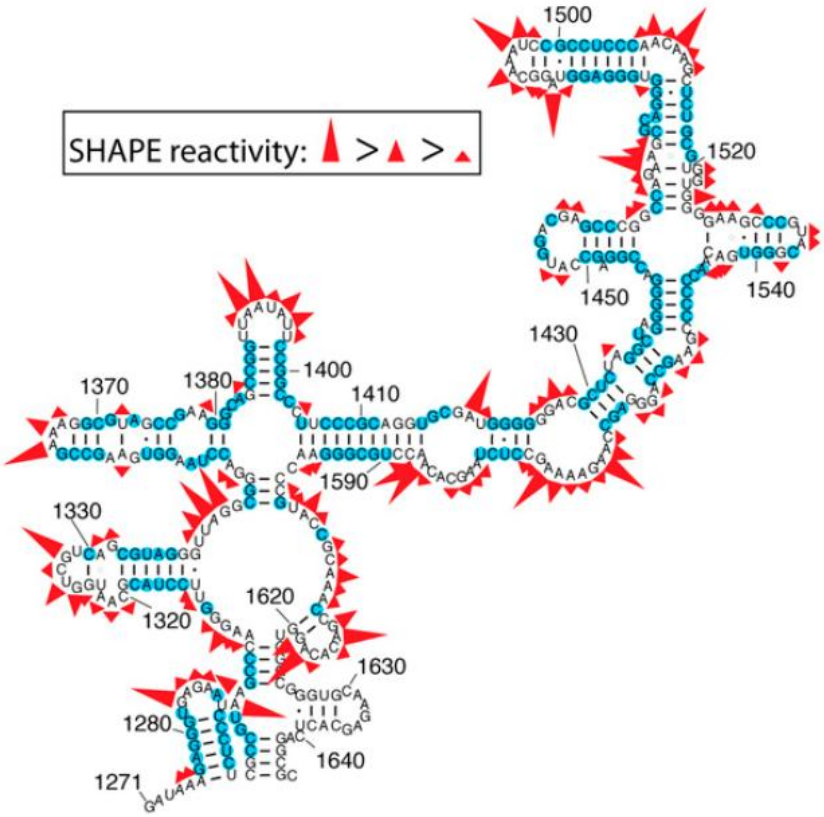


Figure 3.1. SHAPE reactivities for Domain III^{alone} in 250 mM Na^+. The blue nucleotides are unreactive[94].

Folding of Domain III^{alone} to a near-native state requires magnesium ions

Folding of RNAs from secondary structure to their native states, containing long-range tertiary interactions, is known to be generally magnesium-dependent [102,103]. The bacterial and archaeal LSUs are rich in Mg^{2+} ions [104,105] and are stabilized by an RNA-Mg^{2+} motif called the Mg^{2+}-microcluster (Mg-μc) [66,106], defined by: (i) two proximal Mg^{2+} ions chelated by a common bridging phosphate group in the form $Mg_{(a)}$ – O1P – P – O2P – $Mg_{(b)}$; (ii) a 10-membered chelation ring that includes the bridging phosphate in the form $Mg_{(a)}$ – OP – P – O5′ – C5′ – C4′ – C3′ – O3′ – P – OP – $Mg_{(a)}$; (iii) octahedral coordination of both Mg^{2+} ions by a combination of water molecules and RNA phosphate groups; and (iv) nucleotides in non-canonical RNA conformations and unstacked bases.

The native state of Domain III rRNA, as inferred from the 3D structure of the assembled LSU, is stabilized by extensive networks of intradomain tertiary base–base, base–backbone, and backbone-magnesium–backbone interactions (Figure 3.2 A [94]). On the addition of Mg^{2+}, the SHAPE reactivities increase at some sites and decrease at others(Figure 3.2B [94]).

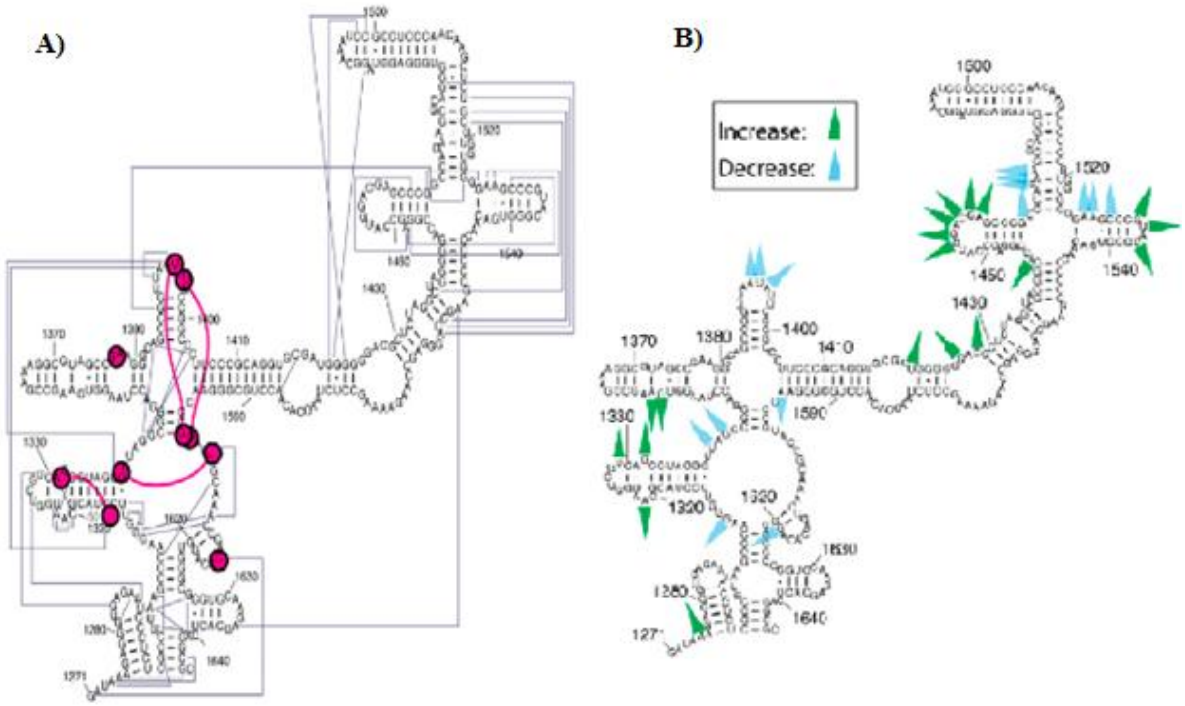


Figure 3.2. (A) Tertiary interactions (dark blue) and phosphate–magnesium–phosphate linkages within Domain III. (B) Magnesium-dependent SHAPE reactivities for Domain III^{alone}, observed upon addition of 10 mM Mg^{2+}.[94]

It was found that in absence of Mg^{2+} the SHAPE reactivities of Domain III^{alone} and Domain III^{23S} are essentially identical along the length of the Domain III sequence as illustrated by the overlaid traces (Figure 3.3A)[94]. The high degree of similarity suggests that the secondary structure of Domain III^{alone} is the same as Domain III^{23S}.Consistent with Mg^{2+}-induced collapse to form a near-native state, the changes in SHAPE reactivity of Domain III upon addition of Mg^{2+} are widely dispersed over Domain III rRNA (Figure 3.3B)[94]. This magnesium dependence of the SHAPE reactivity reflects (i) specific magnesium binding, (ii) more diffuse interactions of magnesium with the RNA, and (iii) tertiary rRNA–rRNA intra-domain interactions. Such magnesium-dependent SHAPE reactivity has previously been demonstrated for tRNA and RNase P[99,107].

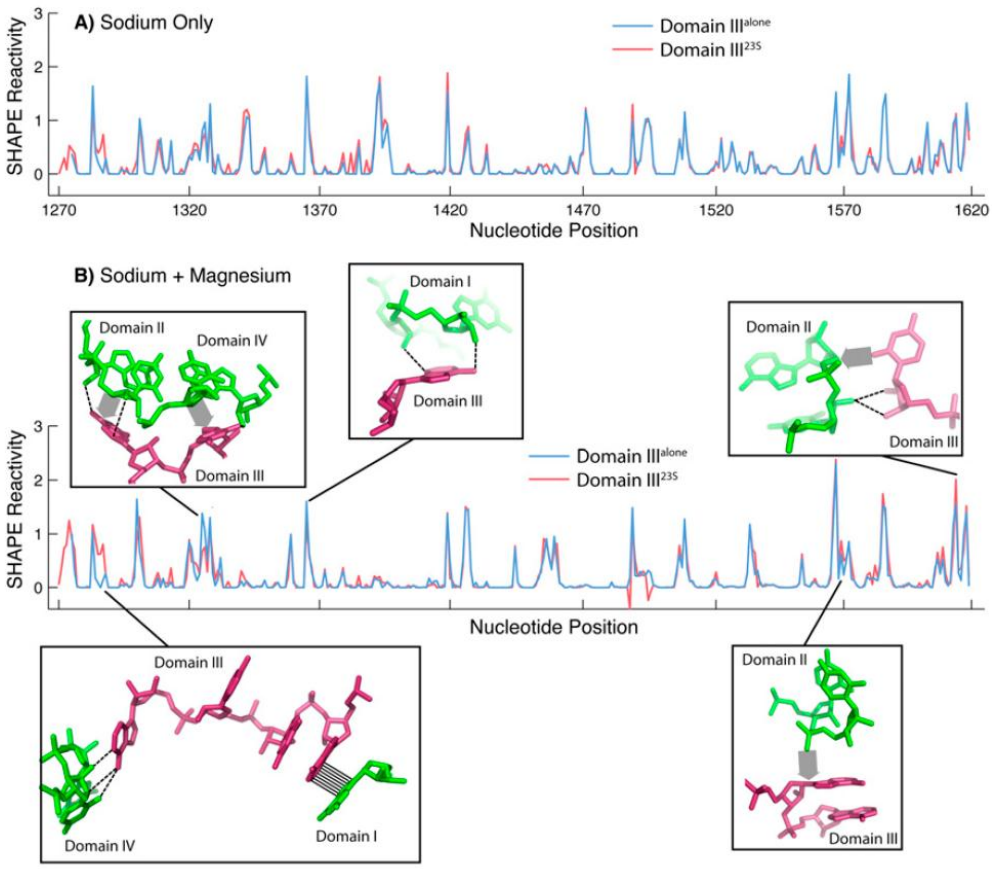


Figure 3.3. SHAPE reactivity for Domain III^{alone} (blue) and Domain III^{23S} (red). (A) Domain III^{alone} and Domain III^{23S} in 250 mM Na^+. (B) Domain III^{alone} and Domain III^{23S} in 250 mM Na^+ and 10 mM Mg^{2+}. Hydrogen bonds are shown by dashed lines, stacking interactions are shown by hashing, and van der Waals contacts are shown by broad shaded arrows[94].

CHAPTER 4

MOLECULAR INTERACTIONS WITHIN DOMAIN III

rRNA-rRNA interactions

The pattern of rRNA tertiary interactions within Domain III of the LSU crystal structure (*T. thermophilus*, PDB entry 2J01 [72]) suggests that Domain III contains an independent subdomain, Domain III^{core}, composed of nucleotides 1271-1406 and 1596-1646. The independent integrity of Domain III^{core} is supported by a large number of intra-subdomain tertiary interactions. Intra-subdomain interactions are both RNA-RNA and RNA-Mg^{2+}-RNA interactions that include base pairing, base stacking, phosphate-base, phosphate-sugar, as well as phosphate-Mg^{2+}-phosphate interactions. Few interactions were observed between Domain III^{core} and Domain III^{H54-59} (Figure 4.1).

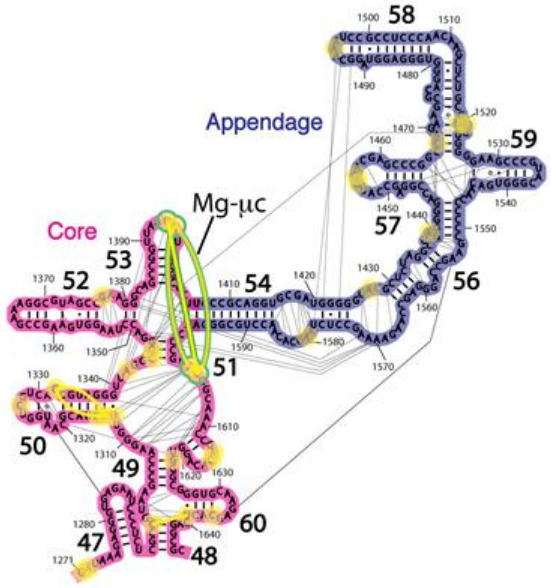


Figure 4.1. Tertiary interactions and phosphate–Mg^{2+}–phosphate linkages within Domain III. Each first shell Mg^{2+}–phosphate interaction is indicated by a yellow circle. The yellow lines between the circles are the phosphate–Mg^{2+}–phosphate linkages.The black lines are rRNA-rRNA interactions as calculated by FR3D [108]

Magnesium interactions

A single Mg-µc is observed in Domain III, which is contained entirely within Domain III^{core}. The two Mg^{2+} ions of this Mg-µc form a tight interaction network with RNA nucleotides of Domain III^{core}, and do not interact with RNA outside of Domain III^{core}. This Mg-µc involves nucleotides A1395, A1603 and C1604, bridging helices 51 and 53 of Domain III^{core} (Figure 4.1).

rRNA-rProtein interactions

Based on the proximity of atoms within the assembled LSU, rProteins L2, L17, L23, and L34 interact with Domain III rRNA. Contacts were defined as a Domain III atom – rProtein atom pair for which the distance between the atoms was less than the sum of their Van der Waals radii. L2, L17, and L34 are bound in rRNA pockets formed partially by Domain III, bridging Domain III and at least two other 2º domains. By contrast, L23 only interacts with Domain I and Domain III (Table 4.1). L23- Domain III contacts are limited to a set of nucleotides localized in Domain III^{core}.

Table 4.1. Interaction[a] of rProteins[b] with 2° Domains of 23S rRNA

Ribosomal protein	I	II	IV	V	VI	III	
						core	appendage
L2	-	+	+	+	-	+	+
L17	-	-	+	-	+	+	+
L23	+	-	-	-	-	+	-
L34	+	+	-	-	-	+	-

[a] rProtein interacting (+) or not interacting (-) with a 2° rRNA domain in the assembled ribosome.
[b]Only rProteins that contact Domain III are listed [PDB 2J01].

Domain III^{core} interactions with L23

The amino acids of L23 that interact with Domain III^{core} are highly conserved over phylogeny. Sequence alignments of L23 on a subset of 121 organisms representing the three domains of life [109] reveal five cationic amino acids K16, K40, K62, K77, and K78 that are highly conserved (present in >90% of the sequences sampled). All conserved amino acids of L23 interact exclusively with Domain III^{core}, suggesting an evolutionary relationship.

Table 4.2. Interactions between Domain III^{core} (DIII^{core}) and highly conserved[a] cationic amino acids of rProtein L23.

Region of L23	L23	DIII^{core}	rRNA moiety	Interaction Type[b,c,d]	Atom Pair	Distance [Å]
Globular Domain	K16	U1340	phosphate	coulombic	NZ-O1P	2.7
	K16	A1393	base	cation-pi	NZ-N7	5.7
	K16	U1394	base	H-bond	NZ-O2	3.0
	K40	A1596	phosphate	coulombic	NZ-O1P	5.0
	K40	A1597	phosphate	coulombic	NZ-O2P	3.9
Extension	K62	U1312	phosphate	coulombic	NZ-O2P	3.3
	K77	U1340	phosphate	coulombic	NZ-O1P	3.7
	K77	U1341	base	cation-pi	NZ-N1	3.3
	K78	U1341	base	H-bond	NZ-O2	2.8

[a]Highly conserved amino acids are those present in >90% of sequences sampled.
[b]Coulombic, interaction between a positively charged NH_3^+ group of a lysine or arginine side chain interact and a negatively charged phosphate groups of the RNA backbone**.**
[c] Hydrogen bonds (H-bonds), interaction between the hydrogen atom of X–H, where X is electronegative, and a basic atom, Y [110].

With geometric analysis, we have parsed the types of molecular interactions (coulombic, cation-pi, and hydrogen bond) that link Domain III^{core} and L23. L23 primarily interacts coulombically with Domain III^{core}: the positively charged NH_3^+ group of a lysine or arginine side chain interacts with a negatively charged phosphate group of

the RNA backbone. Such NH_3^+ -PO_4^- interactions (defined here by a cutoff distance of 5.5 Å) are observed for conserved amino acids K16, K40, K62, and K77 of L23.

Hydrogen bonds (H-bonds) are an attractive interaction between the hydrogen atom of X–H, where X is electronegative, and a basic atom, Y [110]. H-bonds are defined by a cutoff distance of 3.4 Å between atoms X and Y (Table II). L23 interacts with Domain III^{core} bases via H-bonding of K16 to U1340 and K78 to U1341. Two cation-pi interactions are observed between cationic amino acids of L23 and the aromatic systems of bases in Domain III^{core}: K16 with A1393 and K77 with U1341.

Cation-pi interactions [111] occur between the plane of an electron-rich pi system (e.g., nucleotide base) and an adjacent cation (e.g., NH_3^+) [112,113]. The cation-pi interactions are defined by a cutoff distance of 6 Å and a cutoff angle θ of 60^o between the normal to the plane of the pi-system and the vector connecting the center of the pi-system of a nucleobase to the NH_3^+ of an amino acid R-group.

Aims of the thesis

Here we test the hypothesis that Domain III rRNA is comprised of at least one subdomain, which we call Domain III^{core} (Figure 4.2). A subdomain is an autonomously-folding element that is substantially smaller than the full 2° domain [114]. Our subdomain hypothesis was conceived from analysis of molecular interactions within Domain III of the *Thermus thermophilus* crystal structure (PDB 2J01) [72].

We demonstrate experimentally that Domain III can be reduced to Domain III^{core}, which folds and assembles independently of the rest of the 23S rRNA. We constructed a recombinant Domain III^{core} and observe that Domain III^{core} has *in vitro* and *in vivo* function similar to intact Domain III. SHAPE reactivity [115,116] indicates that Domain III^{core} folds in a Mg^{2+}-dependent fashion to the same collapsed state as when it is embedded within Domain III or within the intact 23S rRNA.

Domain III forms part of the exit tunnel in the assembled ribosome. The exit tunnel also includes four evolutionarily conserved ribosomal proteins (L22, L23, L24 and L29) as well as other kingdom-specific proteins. Ribosomal protein L23 is unique because it functions as a docking site for both trigger factor (TF) and signal recognition particle (SRP)[117]. Folded Domain III^{core} appears to contain native-like binding sites for ribosomal proteins and even for a peptide derived from a ribosomal protein. In the LSU, L23 has a globular domain on the LSU surface and an extension (L23^{peptide}) that penetrates into the LSU [118], interacting with the rRNA of DIII^{core}. L23^{peptide} (amino acids 58-79 from *T. thermophilus* L23) traverses the surface of Domain III^{core} (Figure 4.2).

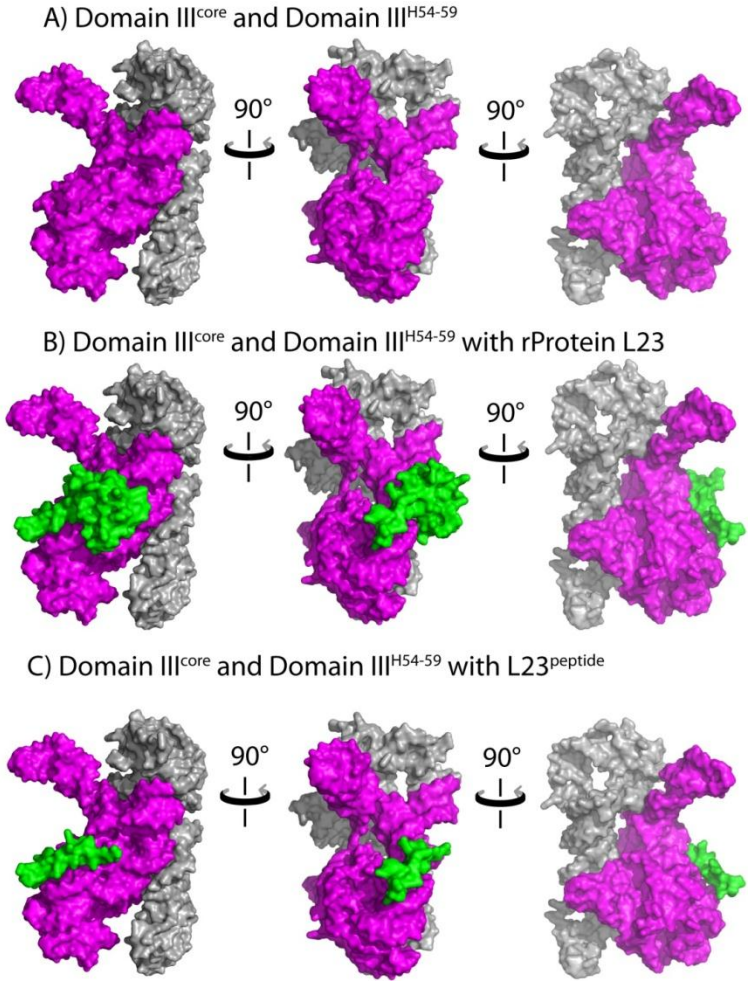


Figure 4.2: A) Domain III in three orientations where DIII^{core} is shown in magenta and DIII^{H54-59} is shown in grey. B) DIII interacts with L23 (green), which is associated with DIII^{core} but not DIII^{H54-59}. C) L23peptide traverses the surface of DIII^{core}.

We analyzed *in vivo* RNA-protein interactions of L23 with DIII and DIII^{core} using the yeast three-hybrid method. Using the continuous variation to analyze *in vitro* RNA-protein interactions, we determined that L23^{peptide} forms a stable 1:1 binary complex with both DIII and DIII^{core}, consistent with interactions observed in the crystal structure of the LSU (Figures 4.2 B & C).

CHAPTER 5

RESULTS

The secondary structure of Domain III^{core} is independently stable

Domain III^{core} is composed of two fragments of rRNA. We joined the two fragments together with a stem-loop. This addition should not perturb the structure since the termini of the two Domain III-core fragments are opposed in a double-stranded A-form helical arm that can directly dock onto the stem-loop (Figure 5.1C). The stem-loop used here (5'-gccGUAAggc-3') is a GNRA stem-loop [119] on a three base pair helical stem (lower case letters).

We used SHAPE [115,116] to probe the structure of DIII^{core}. The SHAPE reactivity of DIII^{core} was initially determined in the absence of Mg^{2+} (Figure 5.1A). Under these conditions, the RNA is expected to form secondary structure but not a collapsed state with long-range tertiary interactions [102,120]. Nucleotides of DIII^{core} were ranked using their absolute SHAPE reactivities relative to A1365 (which had highest SHAPE reactivity). The nucleotides were then binned into four groups, as indicated in Figure 5.1C.

SHAPE reactivity data obtained here for DIII^{core} corresponds very closely with its canonical secondary structure within the 23S rRNA. More importantly, as illustrated by SHAPE, the reactivities of DIII and DIII^{core} are essentially identical along the length of the DIII^{core} sequence, except for the stem-loop connecting the two fragments of DIII^{core}. The high degree of similarity in SHAPE reactivities suggests that the secondary structure of DIII^{core} alone is the same as DIII in the absence of Mg^{2+}.

Mg^{2+}-compacted structure of Domain III^{core} is independently stable

The folding of large RNAs, from secondary structure to collapsed native states with long-range tertiary interactions, is Mg^{2+}-dependent [102,120]. The native state of DIII rRNA, as inferred from the 3D structure of the assembled LSU, is stabilized by extensive networks of intra-domain tertiary base-base, base-backbone and backbone-Mg^{2+}-backbone interactions (including an Mg-μc). Consistent with Mg^{2+}-induced collapse to form a near-native state, the changes in SHAPE reactivity of DIII upon addition of Mg^{2+} are widely dispersed over DIII rRNA (Figure 5.1B).

SHAPE reactivities increase at some sites and decrease at others. This pattern of Mg^{2+}-dependent SHAPE reactivity reflects (i) specific Mg^{2+} binding to RNA, (ii) diffuse interactions of Mg^{2+} with the RNA, and/or (iii) formation of rRNA-rRNA tertiary interactions. Mg^{2+}-dependent SHAPE reactivity has been demonstrated previously for tRNA [116], RNase P [121], and DIII of the LSU [122]. The effect of Mg^{2+} on the SHAPE reactivities of $DIII^{core}$ and DIII is very similar (Figure 5.1B). The Mg2+-dependence of $DIII^{core}$ folding is therefore retained when $DIII^{core}$ is excised from DIII or from the 23S rRNA.

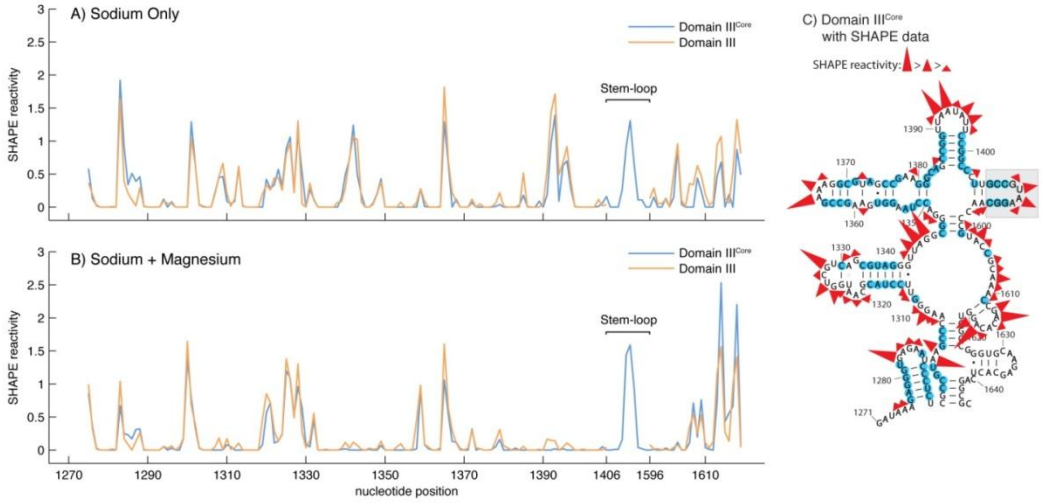


Figure 5.1. Effect of Mg^{2+} on the SHAPE reactivity of $DIII^{core}$ (blue) and DIII (orange). The vertical axis represents SHAPE reactivities and the horizontal axis represents

nucleotide position. A) DIII and DIII^{core} reactivity in 250 mM Na^{+}. B) DIII and DIII^{core} reactivity in 250 mM Na^{+} and 10 mM Mg^{2+} [Courtesy: Dr. Shreyas Athavale].

Some of the most intense Mg^{2+}-induced changes in SHAPE reactivity of DIII^{core} reflect direct Mg^{2+} interactions with the rRNA. The SHAPE reactivity of 20 nucleotides changes by 50% or more upon the addition of Mg^{2+}. Nineteen of these nucleotides interact directly with Mg^{2+} or are in the close proximity to Mg^{2+} in the 3D structure. For example, among nucleotides A1395, A1603 and C1604, which form a Mg^{2+}-μc motif, A1395 and A1603 report changes in SHAPE reactivity of >80% upon addition of magnesium. Other nucleotides directly involved in Mg^{2+} interactions in the crystal structure that give large changes in SHAPE reactivity upon addition of Mg^{2+} are U1313, A1342, A1614 and C1615. Nucleotides G1283, A1284, U1300, U1341, G1343, A1392, A1393, U1396, C1607, A1609, C1617 and A1618 give large Mg^{2+}-induced changes in SHAPE reactivity, and are in close proximity to Mg^{2+} ions.

Domain III and Domain III^{core} interact with L23 and L34 *in vivo*

Based on their association with DIII in the fully assembled ribosome, we predicted DIII and DIII^{core} would interact *in vivo* with rProteins L2, L17, L23 and L34 (Table 5.1). We used the yeast three-hybrid method (Y3H) [123-125] to assay the *in vivo* rRNA-rProtein interactions (Figure 5.2). An RNA-protein interaction in Y3H assay results in resistance to 3-amino-1,2,4-triazole (3-AT) due to increased expression of reporter gene *HIS3* [125].

Table 5.1. Predicted and *in vivo* interaction[a] of rProteins[b] with 2° Domains of 23S rRNA. Interaction of RNA with protein in yeast results in resistance to 3-AT due to increased expression of reporter gene *HIS3*.

Ribosomal protein	**Predicted interaction**		**Yeast three-hybrid assay results**		
	DIII	$DIII^{core}$	DIII	$DIII^{core}$	Resistant up to [3-AT] mM
L2	+	+	-	-	-
L17	+	+	-	-	-
L23	+	+	+	+	**0.3**
L34	+	+	+	+	**<0.1**

[a] rProtein interacting (+) or not interacting (-) with a 2° rRNA domain in the assembled ribosome.

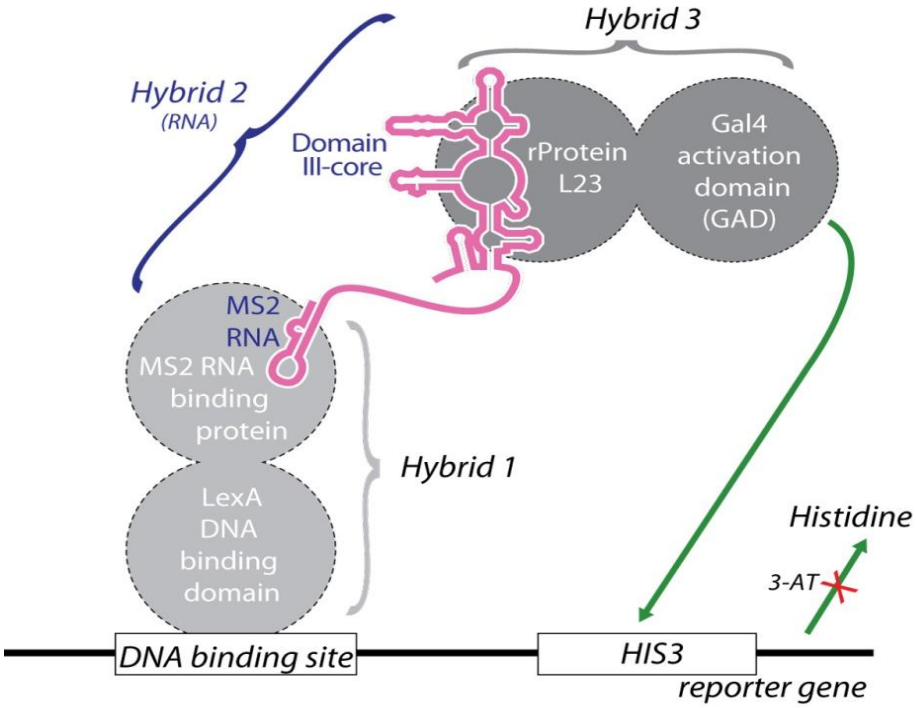
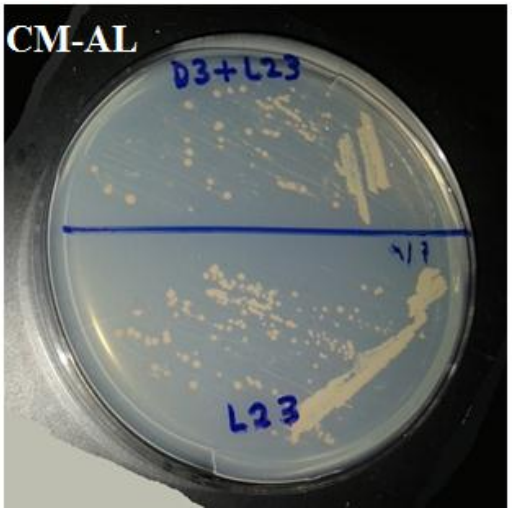


Figure 5.2. Schematic diagram of the yeast-three-hybrid assay for detecting and analyzing RNA-protein interactions. In yeast strain YBZ-1, the *HIS3* reporter gene is under the control of the LexA operator. Hybrid 1, a LexA/MS2 coat protein fusion, binds to the LexA operator. The MS2 coat protein domain binds tightly to the MS2 sequence of the hybrid RNA, which contains the MS2 RNA and RNA sequence of interest, e.g., DIII or DIII^{core}. In Hybrid 3, the protein of interest (e.g., rProtein L23) is fused to the yeast GAL4 transcriptional activation domain (GAD). *In vivo* RNA-protein binding completes Hybrid 2, resulting in expression of the *HIS3* reporter gene. The strength of RNA-protein binding is determined by resistance to 3-Amino-1,2,4-triazole (3-AT), a competitive inhibitor of the *HIS3* product, imidazoleglycerol-phosphate dehydratase [125].

In the Y3H assay, RNA-protein *in vivo* interaction was detected only for p50/p53 (positive control), DIII/L23, DIII^{core}/L23, DIII/L34, DIII^{core}/L34. (Figure 5.3A). *HIS3* reporter gene expression was qualitatively assayed by spreading cell suspensions on plates lacking histidine, adenine (marker for RNA expression vector) and leucine (marker for protein expression vector) (Figure 5.3B).

A)



B)

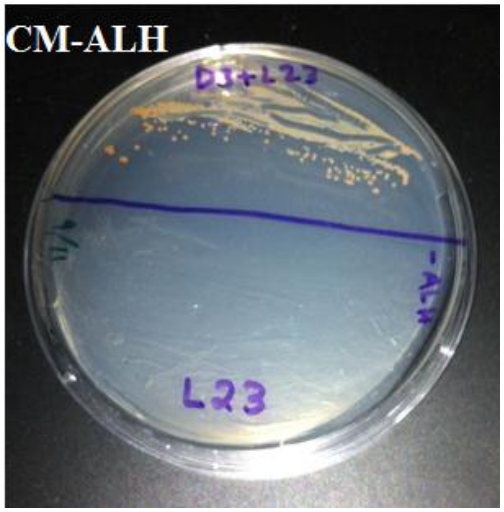
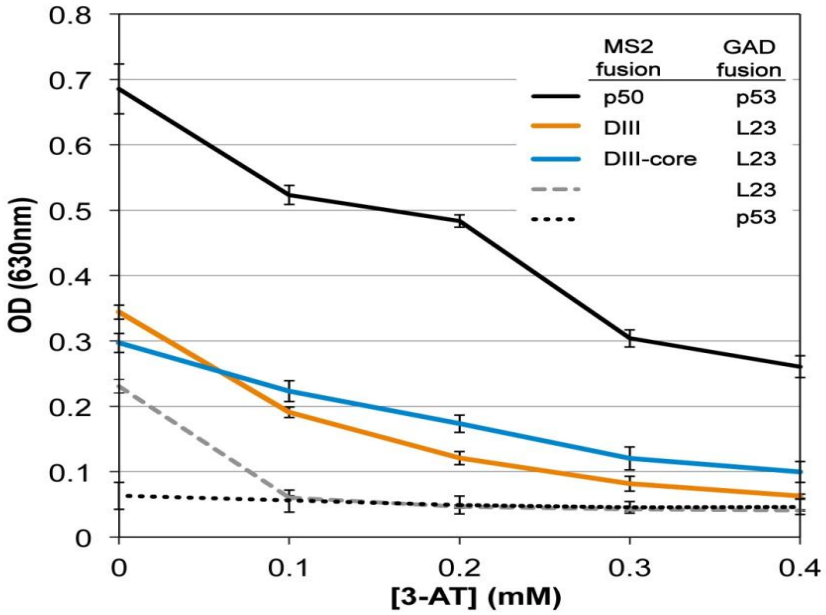


Figure 5.3. GAL4AD-L23 fusion protein expressed and functional in yeast. (A) Growth of yeast cells expressing MS2-DIII and GAD-L23 (D3 + L23); MS2 and GAD-L23 (L23) on plates lacking adenine and leucine (CM-AL). (B) *HIS3* reporter gene expression was assayed by spreading cell suspensions on plates lacking histidine, adenine and leucine and containing histidine biosynthesis inhibitor 3-AT (CM-ALH + 0.2 mM 3-AT).

To quantitatively determine the expression of the *HIS3* reporter gene, the Y3H was assayed at increasing concentrations of 3-AT. *In vivo* interaction of L23 with both DIII and DIII^{core} resulted in similar expression of reporter gene *HIS3*, as determined by a 3-AT resistance of approximately 0.4 mM (Figure 5.4A). *In vivo* interaction of L34 with both DIII and DIII^{core} resulted in similar expression of reporter gene *HIS3*, as determined by a 3-AT resistance of <0.1 mM (Figure 5.4B). Although expression for DIII and DIII^{core} interactions with L23 and L34 is less than that observed for a positive control (p50 RNA/p53 protein), *HIS3* expression is above background (assay without DIII or DIII^{core}), confirming *in vivo* RNA-protein binding.

A)



B)

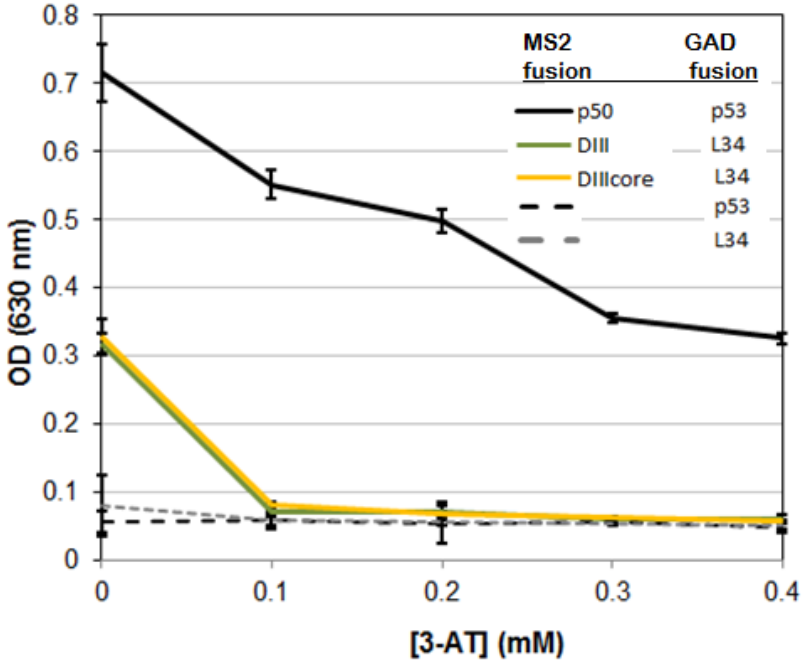
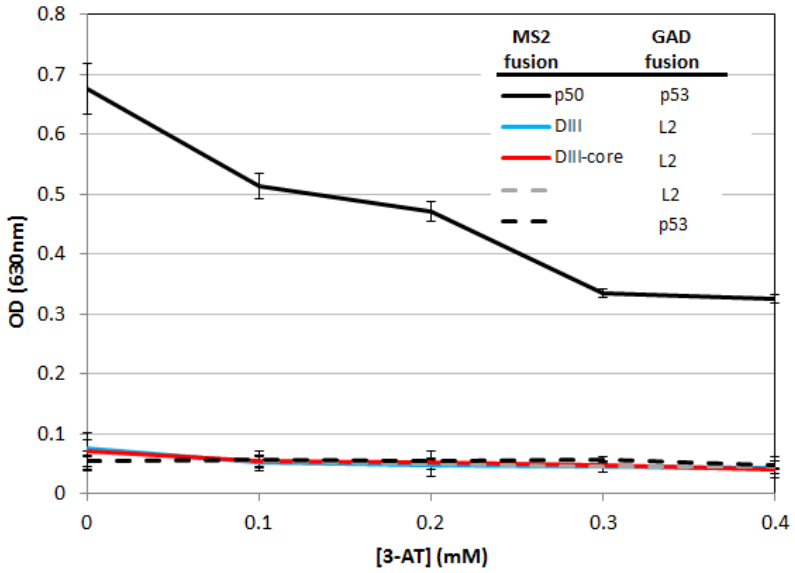


Figure 5.4. Yeast three-hybrid assay results for 3-AT resistance of yeast strain YBZ-1 expressing Domain III-MS2 or DIII^{core} -MS2 and GAD-rProteins (A) rProtein L23 and (B) L34 as well as positive control p50/p53. Cell growth (measured at O.D. 630 nm after 48 hr) is plotted against increasing 3-AT concentration. Negative controls are MS2 RNA alone (lacking DIII or DIII^{core}) assayed with rProteins L23, L34 and p53.

No *in vivo* interactions were observed for rProteins L2 and L17 with either DIII or DIII^{core} (Figures 5.5A and B) though these rProteins are associated with DIII in the fully assembled ribosome.

A)



B)

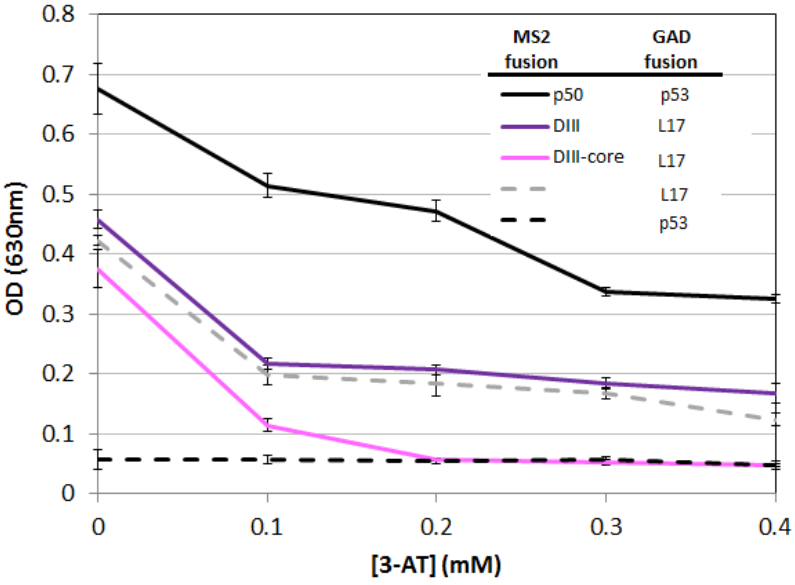


Figure5.5. Yeast three-hybrid assay results for 3-AT resistance of yeast strain YBZ-1 expressing Domain III-MS2 or DIII^{core} -MS2 and rProteins (A) GAD-L2 and (B) GAD-L17 and positive control p50/p53.

L23^{peptide} forms a 1:1 complex with Domain III and with Domain III^{core} in vitro

rProtein L23 has a globular domain on the LSU surface and an extension (L23^{peptide}) that penetrates into the LSU and interacts with the rRNA of Domain III-core. L23^{peptide} penetrates into the LSU and interacts with DIII^{core}. We characterized the binding stoichiometries of the interaction of L23^{peptide} with DIII or DIII^{core} *in vitro* using the method of continuous variation [126,127]. A series of solutions with constant [L23^{peptide}] plus [DIII] or with constant [L23^{peptide}] plus [DIII^{core}] was prepared with variable L23^{peptide}/rRNA ratios. To enable detection by fluorescence, we added an intrinsic fluorophore (Trp) to the C-terminal end of the L23^{peptide}. This amino acid is a tryptophan in the *E. coli* L23, but not in the *T. thermophilus* L23. We measured the increasing fluorescence intensity at 350nm with increasing moles of L23^{peptide} with DIII and DIII^{core} RNA (Figure 5.6).

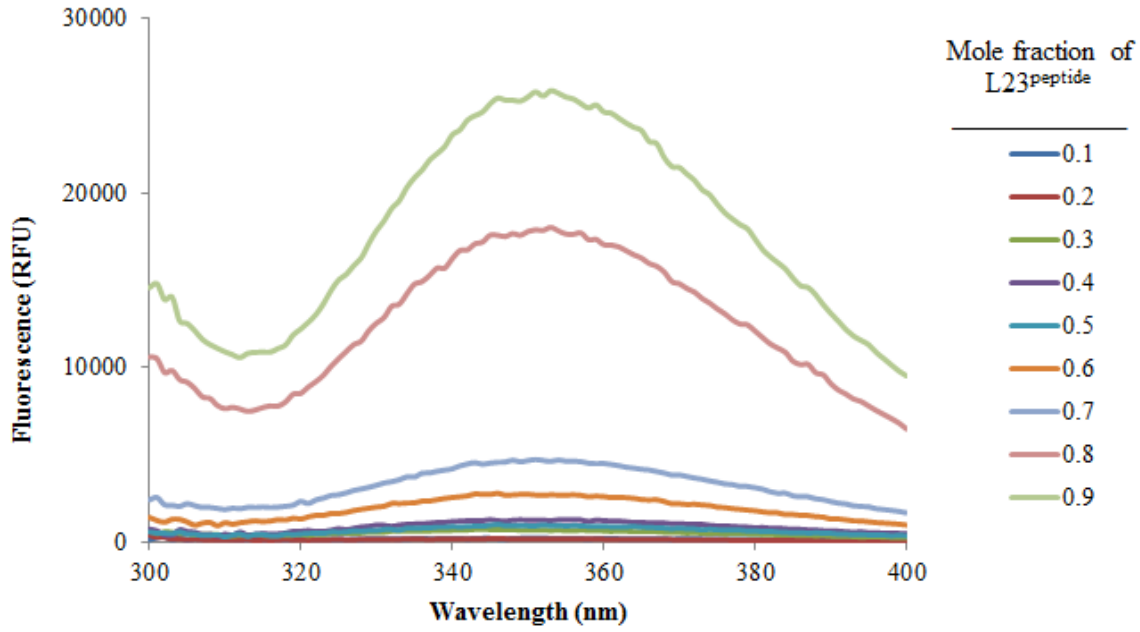


Figure 5.6. Fluorscence spectra of Domain III and L23^{peptide} at different molar ratios. Fluorescent intensity increases with increasing mole fraction of L23^{peptide}.

We inferred the stoichiometry of binding from the discontinuity in the plot of fluorescence intensity versus mole fraction of $L23^{peptide}$. The results indicate that $L23^{peptide}$ forms a 1:1 complex with both DIII and with $DIII^{core}$ (Figure 5.7). These experiments were conducted in the presence of monovalent cations only. Upon the addition of Mg^{2+}, a more complex interaction between $L23^{peptide}$ and both DIII and $DIII^{core}$ is observed (work in progress).

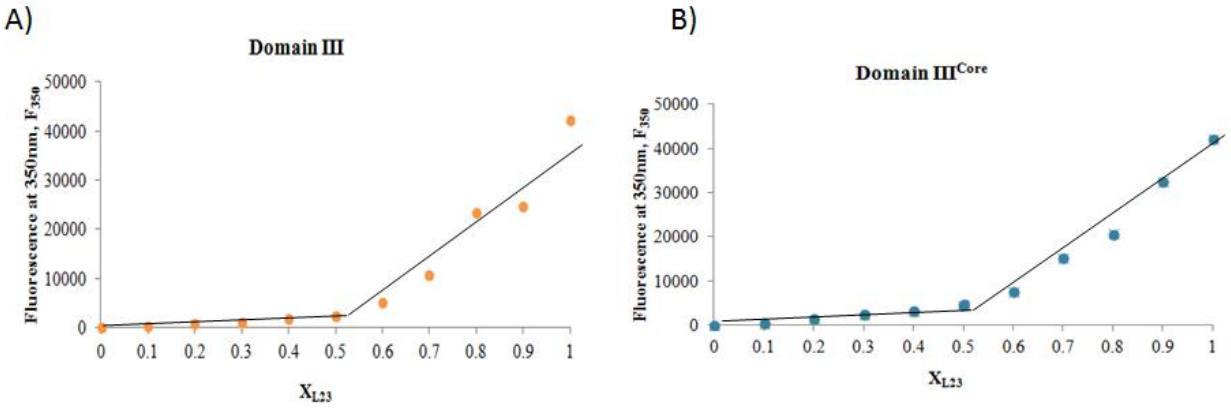


Figure 5.7. Continuous variation of fluorescence intensity versus the mole fraction (X) of L23peptide. A) The second component is DIII rRNA. B) The second component is $DIII^{core}$ rRNA. The trend lines intersect at x=0.53 for DIII and at x=0.52 for $DIII^{core}$, consistent with a 1:1 stoichiometry of interaction of L23peptide with each RNA.

CHAPTER 6

METHODS

Computational analysis of RNA-protein contacts

Coordinates of ribosomal proteins were extracted from the crystal structure of the *T. thermophilus* 70S ribosome (PDB 2J01). The nucleic acid residues of 23S rRNA and amino acid residues of rProteins are named according to the numbering scheme of the *T. thermophilus* crystal structure.

Van der Waals contacts of rProteins with 2° domains of 23S rRNA (Table 4.1) were identified as a DIII – rProtein atom pair for which the distance between the atoms was less than the sum of their Van der Waals radii.

Geometric analysis of molecular interactions

RNA-RNA interactions were taken directly from analysis with FR3D [108] using the same nomenclature as FR3D, which was developed by the BGSU RNA Structural Bioinformatics Lab. Interactions with the neighboring base pair within the same helix were filtered out. The Mg^{2+} - phosphate interactions were calculated directly from the crystal structure: a base-phosphate and Mg^{2+} within 2.6 Å of each other were defined as interacting. RNA-protein calculations were computed directly from the crystal structure: RNA within 3.4 Å of any atom of L23 was defined as contacting. Coulombic interactions were defined at cutoff distance of 5.5 Å between a NH_3^+ of lysine or arginine R-groups and a phosphate group of the RNA backbone**.** H-bonds were defined as a partially positive R-group hydrogen located within 3.4 Å of a partially negative atom of a nucleo base. Cation-pi interactions were defined by a cutoff distance of 6 Å and a cutoff angle θ of 60^o between the normal to the plane of the pi-system and the vector connecting the center of the pi-system of a nucleobase to the NH_3^+ of an amino acid R-group [111].

Construction of transcription vectors for Domain III and Domain III^{core}

DIII rRNA. The DNA gene encoding DIII rRNA was constructed by recursive PCR as described previously [122] and its sequence was confirmed.

DIII^{core} rRNA. DIII^{core} is defined here as 23S rRNA nucleotides G1271-U1406 and A1596-G1647 connected with a 5'gccGUAAggc-3' stem-loop (Figure 5.1C). The numbering scheme of Figure 5.1 is consistent with the numbering of the *T. thermophilus* crystal structure.

The DIII^{core} gene was synthesized in two sequential PCRs with *Pfu* polymerase under standard cycling conditions. In the first PCR, part of DIII^{core} gene, C1295-G1647, was synthesized by recursive PCR [128,129]. In the second PCR, an EcoRI site, T7 promoter (5'-TAATACGACTCACTATAGGG-3'), and the remainder of the DIII^{core} gene (G1271-T1294) was appended upstream, and a HindIII site appended downstream of C1295-G1647.

In the first PCR the forward primer was 5'-CGCCGTAAGCCCAAGGGTT-3', and the reverse primer was 5'-CGCGCCTGAGTGCTCTTGCACC-3'. Four partially complimentary oligos comprised the template: (i) 5'-CGCCGTAAGCCCAAGGGTTCCTACGCAATGGTCGTCAGCGTAGGGTTAGGCGGGACCTAAG-3' (sense), (ii) 5'-TAACCGGCTGCCCTTCGGCTACGCCTTTCGGCTTCACCTTAGGTCCCGCCTAACCCTAC-3' (compliment), (iii) 5'-CGAAGGGCAGCCGGTTAATATTCCGGCCCTTGCCGTAAGGCAACCCGTACCGCAAACCGACACAGGTGGGCG-3' (sense), and (iv) 5'-CGCGCCTGAGTGCTCTTGCACCCGCCCACCTGTGTCGGT-3' (compliment). The stem-loop connecting the two *T. thermophilus* 23S rRNA fragments appears in the fourth oligo (underlined).

In the second PCR the forward primer was 5'-GTGGGAATTCTAATACGACTC-3', and the reverse primer was 5'-CACCAAGCTTCGCGCCTGAGTGCTCTTGCACC-3'. Two partially complementary dsDNAs comprised the template in the second reaction: (i) the purified amplification

product from the first PCR, and (ii) a dsDNA encoding the 5'-EcoRI-T7 promoter-G1271-T1294-3' $DIII^{H54\text{-}H59}$ and a C1295-T1313 complementary region (5'-GTGGGAATTCTAATACGACTCACTATAGGGATAAAGAGGGTGAGAATCCCTC TCGCCGTAAGCCCAAGGGTT-3'). The latter dsDNA was an equimolar mixture of fully complementary, synthetic ssDNAs (MWG Operon).

The product of the second PCR was cloned to pUC19 at the EcoRI and HindIII sites. DH5α transformants positive for insert were sequenced bi-directionally for consensus.

Transcription

Transcription reactions were performed by the run-off method [130], using the MEGAscript High Yield Transcription Kit (Applied Biosystems). pUC19 constructs containing the $DIII^{core}$ sequence were linearized by digestion with HindIII and purified by DNA Clean & Concentrator Kit as described above. Linearized construct (0.5 µg) was transcribed in 20 µL reaction volumes for 14 hours at 37 °C. Transcription reaction conditions were scaled as appropriate to optimize purity and yield. RNA products from transcription reactions were recovered by ammonium acetate precipitation and resuspended in nuclease-free water (IDT). Yields were quantified by UV absorbance.

SHAPE Analysis

Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) methods were adapted from published protocols. SHAPE was performed on DIII and $DIII^{core}$ RNA as described [122]. In vitro-transcribed a-rRNA was prepared in TE buffer (10mM Tris–HCl, 1mM EDTA, pH 8.0) at 100 ng/ml a-rRNA. Thirty-two microliter aliquots of the RNA solution were added to 4 ml of 10X folding buffer (500mM NaHEPES pH 8.0, 2M NaOAc, varying $MgCl_2$) and incubated at 37 °C for 20 min.

SHAPE data were processed as described earlier [122].

Constructs for the yeast three-hybrid assay

L23-GAD protein hybrid. L23 (accession number F6DEQ0) was amplified from genomic DNA (primer sequences in Table III) and cloned into pACTII at restriction sites BamHI and XhoI.

DIII-MS2 and DIII^{core}-MS2 RNA hybrids. From the transcription vectors described above, DIII and DIII^{core} were amplified (primer sequences in Table III) and cloned into the T-cassette vector [131] at the SphI restriction site.

Positive control p50/p53. RNA hybrid p50-MS2 and protein hybrid GAD-p53 vectors were a generous gift from James Maher [131,132].

Yeast three-hybrid assay

Three-hybrid assays were performed in the YBZ-1 yeast strain as described previously [124]. Double transformants were selected in medium lacking adenine and leucine (CM-AL), and subsequently grown in media lacking adenine, leucine and histidine (CM-ALH) with 0, 0.1, 0.2, 0.3, and 0.4 mM 3-AT. Interactions between DIII or DIII^{core} RNA and L23 result in activation of the GAL4 promoter and *HIS3* gene transcription and quantitatively evidenced by growth in CM-ALH media. *HIS3* activity in this assay correlates with RNA–protein affinity as measured *in vitro* over a 10-fold to 100-fold range [124]. Positive controls were p50-MS2 and GAD-p53 [132]. Negative controls included assays of each protein hybrid with only the MS2 RNA (T-cassette vector without insert). Culture growth was determined by optical density (O.D.) at 630nm.

Method of Continuous Variation

The L23^{peptide}, comprising amino acids His58-Ala79 and Trp80, was produced by RS Synthesis (Louisville, KY). Maintaining a constant total concentration of RNA and peptide, the mole fraction of L23^{peptide} to RNA was varied. The concentration of L23^{peptide}

was evaluated after peptide hydrolysis with hydrochloric acid to their constituent amino acids [133]. By measuring the absorbance of Trp80 from the hydrolyzed $L23^{peptide}$, the concentration of $L23^{peptide}$ in the sample was determined to be 110 µM. $L23^{peptide}$ was hydrolyzed in 6 M HCl at 150 °C for 6h in Pierce Vacuum Hydrolysis Tubes. After hydrolysis, sample was concentrated by removing the hydrochloric acid via a stream of Argon directed to the surface of the sample. The dried sample was redissolved in 20 mM Tris-HCl buffer (pH 6.8) and absorbance of Trp80 in the hydrolyzed sample was measured at 278 nm.

Stock solutions (65 µM DIII RNA; 73 µM $DIII^{core}$ RNA; 110 µM $L23^{peptide}$; each in 20 mM Tris-HCL, pH 6.8) were used to make 10 solutions of peptide-RNA in which the mole fraction of peptide was varied from 0.0 to 1.0. The RNA and the peptide were suspended in 10 µL of 20 mM Tris-HCl buffer (pH 8.3) and heated to 85 °C for 30 seconds, then cooled to 30 °C at a rate of 1.5 °C/minute. One microliter of each mixture was placed in a microtiter plate and the fluorescence emission read at 350 nm using Biotek Synergy H4 Multi-Mode Plate Reader.

CHAPTER 7

DISCUSSION

The architecture of the ribosome has profound implications for our understanding of ribosomal assembly and function, and evolution of early life. The architecture of the ribosome can be probed experimentally. Both rRNA and rProteins can be fragmented into hypothetical evolutionary and functional units and the units can be assayed for ability to fold and function.

rRNA is composed, at least in part, of autonomously folding sequences of RNA with intrinsic ability to fold and assemble with proteins and small molecules. These rRNA substructures can be extracted from the ribosome by molecular biological techniques and employed as isolated native-like rRNA components. One of the best-characterized rRNA substructures is Puglisi's Dec-Center$_{27}$ (our nomenclature), a 27-mer RNA extracted from the 16S rRNA [134]. Dec-Center$_{27}$ alone autonomously folds to the conformation observed in small subunit A-site of the intact ribosome. Antibiotics that target the A-site of the small subunit *in vivo* bind *to* Dec-Center$_{27}$ *in vitro*. A second well-characterized rRNA substructure is Draper's GAR$_{58}$ (GTPase-associated region), a 58-mer RNA extracted from the 23S rRNA that autonomously folds to the conformation observed in the large subunit [135]. GAR$_{58}$ rRNA autonomously folds and binds to its ribosomal partner rProtein L11.

Many rProteins have globular domains on the subunit surface, and extensions that penetrate the subunit core [118]. Ribosomal proteins can also be cleaved, by molecular biological techniques, into globular domains and extended tails (rPeptides). In many cases rPeptides bind specifically to their native ribosomal pairing partners. rPeptides derived from rProteins L2, L3, L4, L15, L22 and L23 bind specifically to rRNA substructures.

We have proposed that some of the these extensions of rProteins are molecular fossils that predate the globular protein domain in evolution [18]. In our conception of ancestral ribosomes, small independently-folding RNA elements associated with short peptides. Such complexes assembled to form a primitive peptidyl transferase center, which performed non-coded condensation reactions, without participation of the SSU (also see Fox, [39]).

Here we have deconstructed a portion of the ribosome into relatively small RNA and protein elements that retain the ability to fold and assemble. We have extracted Domain III from the 23S rRNA and further reduced Domain III to Domain III^{core}. Domain III^{core} retains the ability to fold into a near native structure. rProtein L23 specifically interacts *in vivo* with both DIII and DIII^{core}, independently of other components of the ribosome. This interaction suggests that DIII^{core} represents a functional rRNA unit in the context of DIII-L23 assembly. Furthermore, the 'un-structured' extension of rProtein L23 that we call L23^{peptide} sustains the interaction function *in vitro,* with both DIII and DIII^{core}. The L23 extension peptide supports assembly independent of any stabilizing effects contributed by the globular domain of L23. The ability of L23^{peptide} to form a 1:1 complex with both DIII and DIII^{core} suggests that L23^{peptide} is a functional rProtein unit.

The results here are consistent with our model of continuous size distribution of folding and assembly elements. We demonstrate that a short fragment of a ribosomal protein (rProtein extension L23^{peptide}; 21 aa long), does bind specifically to fragments of ribosomal RNA (DIII^{core} and DIII) *in vitro*.

The biological functions of ribosomal protein extensions remain unclear. The extensions of rProteins L4 and L22 can be deleted from *E. coli* without deleterious consequence in ribosomal assembly and function[136]. Mutations in the extensions of L4 and L22 confer macrolide antibiotic resistance in *E. coli* [137]. Little is known about the requirement for the L23 extension, which lines the exit tunnel; however, a recent study suggests that the L23 extension could interact with ribosome-bound nascent polypeptides,

triggering a conformational change in L23 globular domain, where TF binds, to regulate TF recruitment.

Domain III of the 23S is phylogenetically variable and, in some mitochondrial rRNAs, is absent [43]. In fragmented bacterial 23S rRNAs, Domain III contains unique fragmentation site that lacks a stem-loop structure found at other fragmentation sites [138].

rProtein L23 is essential in *E. coli* [139], but expendable in *Bacillus subtilis* [140]. In addition to functions in docking of both TF and SRP, L23 is capable of binding to SecA, which is proposed to post-translationally recognize secretory proteins [141]. L23 is not present in eukaryotic ribosomes; however, rProtein L39e occupies a similar position in the eukaryotic and archaeal exit tunnel [142].

The ribosome as whole is quite robust and has functional plasticity. L23 is multi-functional yet is not universally necessary for function; similarly, Domain III is highly variable and not universally required for translation. It is intriguing that this dynamic ribosomal protein maintains *in vivo* binding to DIII rRNA in the absence of supporting interactions from other rRNA and rProteins in the assembled ribosome. Even further lacking supporting interactions; $L23^{peptide}$ maintains 1:1 *in vitro* binding stoichiometry with $DIII^{core}$.

Our work supports a platform on which deconstructing the ribosome and exploring its units of modification further elucidates structure, function and evolution of this complex molecular machine. In addition, the non-uniform distribution of the DIII-L23 interaction unit offers targets for antibiotics. The ability to isolate the DIII-L23 interaction from the fully assembled ribosome both *in vivo* and *in vitro* also offers relatively simple, and potentially high-throughput, assays to test antibiotic activity.

Evolutionary implications of ribosomal deconstruction

Structure, like sequence, reveals information about macromolecular origins and evolution. The pioneering work on structural elucidation of the ribosome [16,67,81] allows a detailed investigation of ribosomal evolution. The studies presented here have demonstrated that the modern ribosome holds fossilized motifs that can be traced back to decipher the origins of ribosome. Deconstruction allows us to look at the ribosome in pieces, rather than a whole, and compare the pieces relative to one another. A skeletal form of the ribosome is most likely to comprise of structural and functional elements that were integral to the earliest form of ribosome, as it would have existed in early life.

Reducing down the ribosome to its minimal components able to retain its most essential function, peptidyl transferase activity, was first attempted in 1982 [143]. Since then, numerous models have been proposed about the earliest forms of ribosome including the concept of proto-ribosome. This ancient ribozyme was a dimeric RNA assembly (Figure 7.1) postulated to have formed spontaneously by gene duplication or gene fusion. The proto-ribosome is posited to have catalyzed non-coded peptide bond formation and elongation [16].

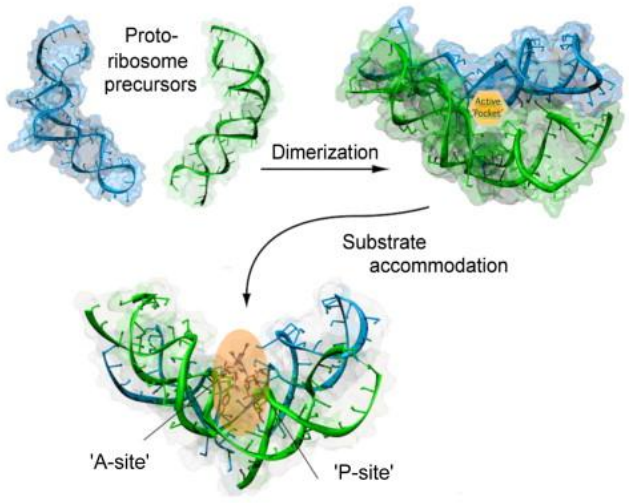


Figure 7.1. Schematic representation of ‘pocket-like’ proto-ribosome showing simple catalytic peptidyl transferase activity [Adapted from [16]].

Relics of the proto-ribosome can be found in the symmetrical rRNA pocket located at the active site of the modern ribosome [16,52,144]. In an otherwise asymmetric ribosome, this symmetric "pocket" is accommodated in the PTC and remains highly conserved (Figure 7.2).

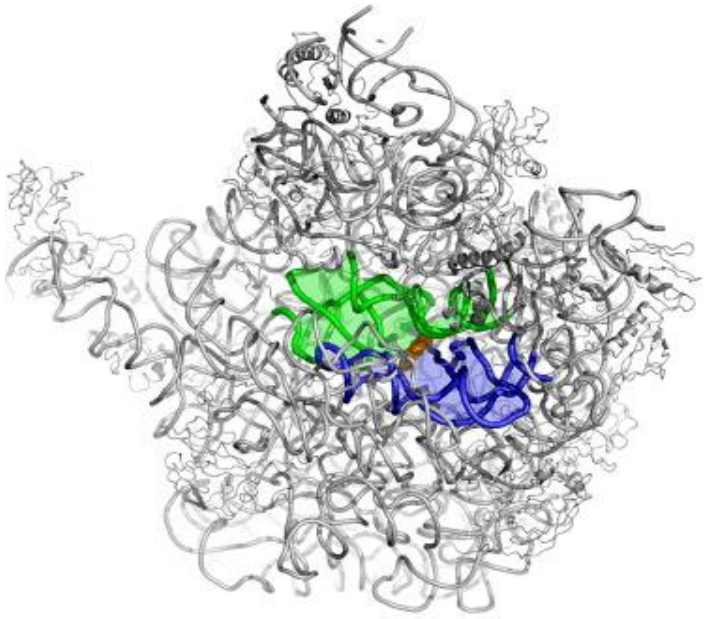


Figure 7.2. Symmetrical region within the large ribosomal subunit. The remnants of the proto-ribosome is shown in blue and green [Adapted from [16]].

Recent developments and future directions

Recently, a model of ancestral PTC (a-PTC) designed from a consensus among previous proposals of ribosomal evolution has shown experimentally that fragments of rRNA and protein from the oldest part of the extant ribosome (Figure 7.3) can associate both *in vitro* and *in vivo*. The rRNA fragments were joined together to form a single RNA polymer called the a-rRNA which folds to a near-native state independently of the remainder of the LSU [45].

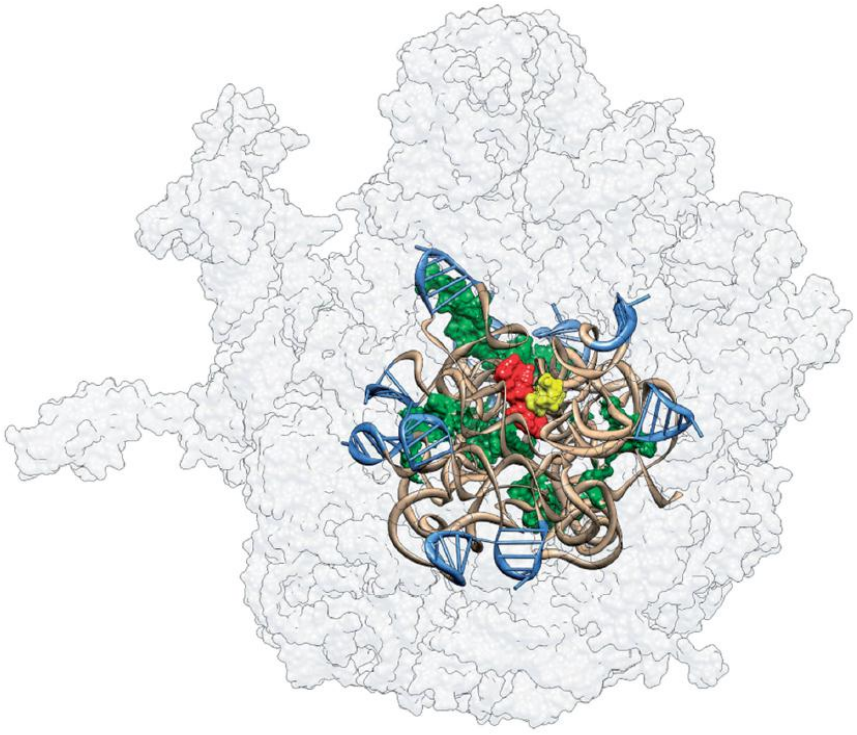


Figure 7.3. 3D model of the a-PTC within the modern LSU surface (light gray, transparent). a-rRNA (shown in brown ribbon and blue stem loops) is shown in contact with five ancestral fragments of ribosomal proteins L2, L3, L4, L15 and L22 (in surface representation green). [Adapted from [45]]

The model of the a-PTC contains the proto-ribosome. Unlike the proto-ribosome, which is devoid of peptides, the a-PTC contains fragments of rProteins that interact specifically with the rRNA of a-PTC. The proto-ribosome appears to represent a model of a smaller, even more primitive ancestor of the a-PTC [45].

Though both a-PTC and Domain III demonstrate independent folding and assembly, Domain III is not included in the model of a-PTC. Parts of Domain V of the 23S rRNA are believed to be among the most ancient parts of the ribosome [10,17,48,145] while Domain III is thought to be a more recent addition [41]. Earlier studies [94] and the results presented in this thesis support the hypothesis that Domain III was added as an intact entity to the ancestral ribosome. Such an evolutionary model is consistent with the absence of Domain III from certain mitochondrial rRNAs, such as that of *Trypanosoma brucei* [146].

In our conception of ribosomal origins, small independently-folding RNA elements associated with short peptides. The PTC evolved into the modern LSU, in a series of cooptions that left unaltered the basic structure and function of the PTC. This model predicts a continuous size distribution of folding and assembly elements within the LSU. We anticipate autonomy and specificity of folding and interaction of small, mid-sized and large rRNA and protein components.

For future work, we would continue to probe ribosomal assembly to identify other ribosomal fragments that are capable of folding and assembly, independent of the remainder of the ribosome. We propose to develop a simple assay to screen for small molecules that can disrupt the interaction between such fragments of rRNA and rProteins much like as the interaction between Domain III and $L23^{peptide}$. Such a screening method will help in the identification of potential lead compounds for development of antibiotics targeting ribosomal assembly in pathogens.

REFERENCES

1. Scheper GC, van der Knaap MS, Proud CG (2007) Translation matters: protein synthesis defects in inherited disease. Nature Reviews Genetics 8: 711-723.
2. Stelzl U, Connell S, Nierhaus K, Wittmann-Liebold B (2001) Ribosomal proteins: role in ribosomal functions: John Wiley & Sons.
3. Noller HF (2004) The driving force for molecular evolution of translation. RNA 10: 1833-1837.
4. Woese C (1967) The genetic code.
5. Crick FH (1968) The origin of the genetic code. J Mol Biol 38: 367-379.
6. Jeffares DC, Poole AM, Penny D (1998) Relics from the RNA World. J Mol Evol 46: 18-36.
7. Poole A, Jeffares D, Penny D (1999) Early evolution: prokaryotes, the new kids on the block. BioEssays 21: 880-889.
8. Woese C (1970) Molecular mechanics of translation: a reciprocating ratchet mechanism. Nature 226: 817-820.
9. Crick FH, Brenner S, Klug A, Pieczenik G (1976) A speculation on the origin of protein synthesis. Orig Life 7: 389-397.
10. Woese CR (2001) Translation: in retrospect and prospect. RNA 7: 1055-1067.
11. Gesteland RF, Cech T, Atkins JF (2006) The RNA world: the nature of modern RNA suggests a prebiotic RNA world: Cold Spring Harbor Laboratory Pr.
12. Egel R (2009) Peptide-dominated membranes preceding the genetic takeover by RNA: latest thinking on a classic controversy. BioEssays 31: 1100-1109.
13. Kauffman SA (1993) The origins of order: Self-organization and selection in evolution: Oxford University Press, USA.
14. Caetano-Anollés D, Kim KM, Mittenthal JE, Caetano-Anollés G (2011) Proteome evolution and the metabolic origins of translation and cellular life. Journal of molecular evolution 72: 14-33.
15. Caetano-Anollés G, Wang M, Caetano-Anolles D, Mittenthal J (2009) The origin, evolution and structure of the protein world. Biochem J 417: 621-637.
16. Davidovich C, Belousoff M, Bashan A, Yonath A (2009) The evolving ribosome: from non-coded peptide bond formation to sophisticated translation machinery. Res Microbiol 160: 487-492.
17. Bokov K, Steinberg SV (2009) A hierarchical model for evolution of 23S ribosomal RNA. Nature 457: 977-980.
18. Hsiao C, Mohan S, Kalahar BK, Williams LD (2009) Peeling the onion: ribosomes are ancient molecular fossils. Mol Biol Evol 26: 2415-2425.
19. Goodwin JT, Mehta AK, Lynn DG (2012) Digital and Analog Chemical Evolution. Accounts of Chemical Research 45: 2189-2199.
20. Cech TR (2012) The RNA worlds in context. Cold Spring Harb Perspect Biol 4: a006742.
21. Roberts E, Sethi A, Montoya J, Woese CR, Luthey-Schulten Z (2008) Molecular signatures of ribosomal evolution. Proceedings of the National Academy of Sciences 105: 13953-13958.

22. Woese C (1998) The universal ancestor. Proceedings of the National Academy of Sciences 95: 6854-6859.
23. Doolittle WF (1999) Phylogenetic Classification and the Universal Tree. Science 284: 2124-2128.
24. Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci U S A 93: 10268-10273.
25. Kyrpides N, Overbeek R, Ouzounis C (1999) Universal protein families and the functional content of the last universal common ancestor. J Mol Evol 49: 413-423.
26. Ribas de Pouplana L, Fox GE, Naik AK (2005) The Evolutionary History of the Translation Machinery
The Genetic Code and the Origin of Life. Springer US. pp. 92-105.
27. Williamson JR (2009) The ribosome at atomic resolution. Cell 139: 1041-1043.
28. Nomura M, Traub P (1968) Structure and function of Escherichia coli ribosomes. 3. Stoichiometry and rate of the reconstitution of ribosomes from subribosomal particles and split proteins. J Mol Biol 34: 609-619.
29. Traub P, Nomura M (1968) Structure and function of Escherichia coli ribosomes. I. Partial fractionation of the functionally active ribosomal proteins and reconstitution of artificial subribosomal particles. J Mol Biol 34: 575-593.
30. Traub P, Nomura M (1968) Structure and function of E. coli ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. Proc Natl Acad Sci U S A 59: 777-784.
31. Nierhaus KH, Dohme F (1974) Total reconstitution of functionally active 50S ribosomal subunits from Escherichia coli. Proc Natl Acad Sci U S A 71: 4713-4717.
32. Rohl R, Nierhaus KH (1982) Assembly map of the large subunit (50S) of Escherichia coli ribosomes. Proc Natl Acad Sci U S A 79: 729-733.
33. Held WA, Mizushima S, Nomura M (1973) Reconstitution of Escherichia coli 30 S ribosomal subunits from purified molecular components. J Biol Chem 248: 5720-5730.
34. Nomura M, Erdmann VA (1970) Reconstitution of 50S ribosomal subunits from dissociated molecular components. Nature 228: 744-748.
35. Rohl R, Nierhaus KH (1982) Assembly map of the large subunit (50S) of *Escherichia coli* ribosomes. Proc Natl Acad Sci U S A 79: 729-733.
36. Fox GE, Naik AK (2005) The evolutionary history of the translation machinery. The genetic code and the origin of life: 92-105.
37. Traub P, Nomura M (1968) Structure and function of E. coli ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. Proceedings of the National Academy of Sciences of the United States of America 59: 777.
38. Clark CG (1987) On the evolution of ribosomal RNA. Journal of molecular evolution 25: 343-350.
39. Fox GE (2010) Origin and evolution of the ribosome. Cold Spring Harb Perspect Biol 2: a003483.

40. Mears JA, Cannone JJ, Stagg SM, Gutell RR, Agrawal RK, et al. (2002) Modeling a minimal ribosome based on comparative sequence analysis. J Mol Biol 321: 215-234.
41. Hury J, Nagaswamy U, Larios-Sanz M, Fox GE (2006) Ribosome origins: the relative age of 23S rRNA domains. Origins of Life and Evolution of Biospheres 36: 421-429.
42. Smit S, Widmann J, Knight R (2007) Evolutionary rates vary among rRNA structural elements. Nucleic Acids Res 35: 3339-3354.
43. Mears JA, Cannone JJ, Stagg SM, Gutell RR, Agrawal RK, et al. (2002) Modeling a minimal ribosome based on comparative sequence analysis. Journal of Molecular Biology 321: 215-234.
44. Hury J, Nagaswamy U, Larios-Sanz M, Fox GE (2006) Ribosome origins: the relative age of 23S rRNA Domains. Orig Life Evol Biosph 36: 421-429.
45. Hsiao C, Lenz TK, Peters JK, Fang PY, Schneider DM, et al. (2013) Molecular paleontology: a biochemical model of the ancestral ribosome. Nucleic Acids Res.
46. Woese CR (2000) Interpreting the universal phylogenetic tree. Proc Natl Acad Sci U S A 97: 8392-8396.
47. Williams D, Fournier GP, Lapierre P, Swithers KS, Green AG, et al. (2011) A rooted net of life. Biology direct 6: 45.
48. Wolf YI, Koonin EV (2007) On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. Biol Direct 2: 14.
49. Gutell RR, Larsen N, Woese CR (1994) Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol Rev 58: 10-26.
50. Doudna JA, Rath VL (2002) Structure and Function of the Eukaryotic Ribosome-The Next Frontier. Cell 109: 153-156.
51. Polacek N, Mankin AS (2005) The ribosomal peptidyl transferase center: structure, function, evolution, inhibition. Critical reviews in biochemistry and molecular biology 40: 285-311.
52. Belousoff MJ, Davidovich C, Zimmerman E, Caspi Y, Wekselman I, et al. (2010) Ancient machinery embedded in the contemporary ribosome. Biochem Soc Trans 38: 422-427.
53. Mueller F, Sommer I, Baranov P, Matadeen R, Stoldt M, et al. (2000) The 3D arrangement of the 23 S and 5 S rRNA in the Escherichia coli 50 S ribosomal subunit based on a cryo-electron microscopic reconstruction at 7.5 A resolution. J Mol Biol 298: 35-59.
54. Harish A, Caetano-Anolles G (2012) Ribosomal history reveals origins of modern protein synthesis. PLoS One 7: e32776.
55. Shajani Z, Sykes MT, Williamson JR (2011) Assembly of bacterial ribosomes. Annual review of biochemistry 80: 501-526.
56. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy IV FV, et al. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science 326: 688-694.
57. Agmon I, Auerbach T, Baram D, Bartels H, Bashan A, et al. (2003) On peptide bond formation, translocation, nascent protein progression and the regulatory properties of ribosomes. European Journal of Biochemistry 270: 2543-2556.

58. Fox GE, Woese CR (1975) The architecture of 5S rRNA and its relation to function. J Mol Evol 6: 61-76.
59. Tal M (1969) Metal ions and ribosomal conformation. Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis 195: 76-86.
60. Bowman JC, Lenz TK, Hud NV, Williams LD (2012) Cations in charge: magnesium ions in RNA folding and catalysis. Current opinion in structural biology.
61. Norris JS, Clawson GA, Schmidt MG, Hoel B, Pan WH, et al. (2010) Tissue-specific and target RNA-specific ribozymes. Google Patents.
62. Holm N. Oceanic lithosphere, Mg (II), Na (I), PPi and life's origin; 2012. Deep Carbon Observatory.
63. Klein DJ, Moore PB, Steitz TA (2004) The Roles of Ribosomal Proteins in the Structure Assembly, and Evolution of the Large Ribosomal Subunit. Journal of Molecular Biology 340: 141-177.
64. Cate JH, Hanna RL, Doudna JA (1997) A magnesium ion core at the heart of a ribozyme domain. Nat Struct Biol 4: 553-558.
65. Juneau K, Podell E, Harrington DJ, Cech TR (2001) Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA–solvent interactions. Structure 9: 221-231.
66. Petrov AS, Bowman JC, Harvey SC, Williams LD (2011) Bidentate RNA-magnesium clamps: on the origin of the special role of magnesium in RNA folding. RNA 17: 291-297.
67. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 289: 905-920.
68. Ramakrishnan V. The ribosome: some hard facts about its structure and hot air about its evolution; 2009. Cold Spring Harbor Laboratory Press.
69. Noller HF (2012) Evolution of protein synthesis from an RNA world. Cold Spring Harbor Perspectives in Biology 4.
70. Moore PB, Steitz TA (2011) The roles of RNA in the synthesis of protein. Cold Spring Harbor Perspectives in Biology 3.
71. Noller HF, Hoffarth V, Zimniak L (1992) Unusual resistance of peptidyl transferase to protein extraction procedures. Science 256: 1416-1419.
72. Selmer M, Dunham CM, Murphy FV, Weixlbaumer A, Petry S, et al. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. Science 313: 1935-1942.
73. Zirbel CL, Šponer JE, Šponer J, Stombaugh J, Leontis NB (2009) Classification and energetics of the base-phosphate interactions in RNA. Nucleic Acids Res 37: 4898-4918.
74. Apostolico A, Ciriello G, Guerra C, Heitsch CE, Hsiao C, et al. (2009) Finding 3D motifs in ribosomal RNA structures. Nucleic Acids Res 37: e29-e29.
75. Laing C, Schlick T (2010) Computational approaches to 3D modeling of RNA. Journal of Physics: Condensed Matter 22: 283101.
76. Wang D, Severinov K, Landick R (1997) Preferential interaction of the his pause RNA hairpin with RNA polymerase β subunit residues 904–950 correlates with strong transcriptional pausing. Proceedings of the National Academy of Sciences 94: 8433-8438.

77. Cruz JA, Westhof E (2011) Sequence-based identification of 3D structural modules in RNA with RMDetect. Nature Methods 8: 513-519.
78. Schimmel P (2008) Development of tRNA synthetases and connection to genetic code and disease. Protein Sci 17: 1643-1652.
79. Gupta A, Gribskov M (2011) The role of RNA sequence and structure in RNA–protein interactions. Journal of Molecular Biology 409: 574-587.
80. Giovanni C, Claudio G, Concettina G (2010) Analysis of interactions between ribosomal proteins and RNA structural motifs. BMC Bioinformatics 11.
81. Selmer M, Dunham CM, Murphy FVt, Weixlbaumer A, Petry S, et al. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. Science 313: 1935-1942.
82. Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, et al. (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. Cell 102: 615-623.
83. Weitzmann CJ, Cunningham PR, Nurse K, Ofengand J (1993) Chemical evidence for domain assembly of the *Escherichia coli* 30S ribosome. FASEB J 7: 177-180.
84. Samaha RR, O'Brien B, O'Brien TW, Noller HF (1994) Independent in vitro assembly of a ribonucleoprotein particle containing the 3' domain of 16S rRNA. Proc Natl Acad Sci U S A 91: 7884-7888.
85. Agalarov SC, Selivanova OM, Zheleznyakova EN, Zheleznaya LA, Matvienko NI, et al. (1999) Independent *in vitro* assembly of all three major morphological parts of the 30S ribosomal subunit of *Thermus thermophilus*. Eur J Biochem 266: 533-537.
86. Brimacombe R, Maly P, Zwieb C (1983) The structure of ribosomal RNA and its organization relative to ribosomal protein. Prog Nucleic Acid Res Mol Biol 28: 1-48.
87. Wimberly BT, Brodersen DE, Clemons WM, Jr., Morgan-Warren RJ, Carter AP, et al. (2000) Structure of the 30S ribosomal subunit. Nature 407: 327-339.
88. Noller HF (2005) RNA structure: reading the ribosome. Science 309: 1508-1514.
89. Powers T, Noller HF (1995) Hydroxyl radical footprinting of ribosomal proteins on 16S rRNA. RNA 1: 194-209.
90. Stern S, Powers T, Changchien LM, Noller HF (1989) RNA-protein interactions in 30S ribosomal subunits: folding and function of 16S rRNA. Science (New York, NY) 244: 783-790.
91. Branlant C, Krol A, Sriwidada J, Brimacombe R (1976) RNA sequences associated with proteins L1, L9, and L5, L18, L25, in ribonucleoprotein fragments isolated from the 50-S subunit of Escherichia coli ribosomes. Eur J Biochem 70: 483-492.
92. Ehresmann C, Stiegler P, Carbon P, Ungewickell E, Garrett RA (1980) The topography of the 5' end of 16-S RNA in the presence and absence of ribosomal proteins S4 and S20. Eur J Biochem 103: 439-446.
93. Noller HF (1984) Structure of ribosomal RNA. Annu Rev Biochem 53: 119-162.
94. Athavale SS, Gossett JJ, Hsiao C, Bowman JC, O'Neill E, et al. (2012) Domain III of the T. thermophilus 23S rRNA folds independently to a near-native state. RNA 18: 752-758.

95. Glotz C, Zwieb C, Brimacombe R, Edwards K, Kossel H (1981) Secondary structure of the large subunit ribosomal RNA from Escherichia coli, Zea mays chloroplast, and human and mouse mitochondrial ribosomes. Nucleic Acids Res 9: 3287-3306.
96. Leffers H, Kjems J, Ostergaard L, Larsen N, Garrett RA (1987) Evolutionary relationships amongst archaebacteria. A comparative study of 23 S ribosomal RNAs of a sulphur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. J Mol Biol 195: 43-61.
97. Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, et al. (2001) Crystal structure of the ribosome at 5.5 A resolution. Science 292: 883-896.
98. Wilkinson KA, Merino EJ, Weeks KM (2005) RNA SHAPE Chemistry Reveals Nonhierarchical Interactions Dominate Equilibrium Structural Transitions in tRNAAsp Transcripts. Journal of the American Chemical Society 127: 4659-4667.
99. Merino EJ, Wilkinson KA, Coughlan JL, Weeks KM (2005) RNA Structure Analysis at Single Nucleotide Resolution by Selective 2‘-Hydroxyl Acylation and Primer Extension (SHAPE). Journal of the American Chemical Society 127: 4223-4231.
100. Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, et al. (2002) The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics 3.
101. Deigan KE, Li TW, Mathews DH, Weeks KM (2009) Accurate SHAPE-directed RNA structure determination. Proc Natl Acad Sci U S A 106: 97-102.
102. Draper DE (2008) RNA folding: thermodynamic and molecular descriptions of the roles of ions. Biophysical Journal 95: 5489-5495.
103. Brion P, Westhof E (1997) Hierarchy and dynamics of RNA folding. Annu Rev Biophys Biomol Struct 26: 113-137.
104. Klein DJ, Moore PB, Steitz TA (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit. RNA 10: 1366-1379.
105. Hsiao C, Tannenbaum M, VanDeusen H, Hershkovitz E, Perng G, et al. (2008) Complexes of Nucleic Acids with Group I and II Cations. In: Hud N, editor. Nucleic Acid Metal Ion Interactions. London: The Royal Society of Chemistry. pp. 1-35.
106. Hsiao C, Williams LD (2009) A recurrent magnesium-binding motif provides a framework for the ribosomal peptidyl transferase center. Nucleic Acids Res 37: 3134-3142.
107. Mortimer SA, Weeks KM (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. J Am Chem Soc 129: 4144-4145.
108. Sarver M, Zirbel CL, Stombaugh J, Mokdad A, Leontis NB (2008) FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. J Math Biol 56: 215-252.
109. Petrov A, Bernier C, Hsiao C, Okafor CD, Tannenbaum E, et al. (2012) RNA-Magnesium-Protein Interactions in Large Ribosomal Subunit. Journal of Physical Chemistry B 116: 8113-8120.
110. Jeffrey GA (1997) An Introduction to Hydrogen Bonding. New York: Oxford University Press.

111. Gallivan JP, Dougherty DA (1999) Cation-pi interactions in structural biology. Proc Natl Acad Sci U S A 96: 9459-9464.
112. Dougherty DA (1997) Electrostatic models for the cation-pi interaction and related non-covalent interactions. J Am Chem Soc 214: 166.
113. Ma JC, Dougherty DA (1997) The cation-pi interaction. Chem Rev 97: 1303-1324.
114. Ecker DJ, Griffey RH (1999) RNA as a small-molecule drug target: doubling the value of genomics. Drug Discov Today 4: 420-429.
115. Merino EJ, Wilkinson KA, Coughlan JL, Weeks KM (2005) RNA Structure Analysis at Single Nucleotide Resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). J Am Chem Soc 127: 4223-4231.
116. Wilkinson KA, Merino EJ, Weeks KM (2005) RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in tRNA(Asp) transcripts. J Am Chem Soc 127: 4659-4667.
117. Baneyx F (2012) A ribosomal surprise. Biotechnology Journal 7: 326-327.
118. Wilson DN, Nierhaus KH (2005) Ribosomal Proteins in the Spotlight. Critical Reviews in Biochemistry and Molecular Biology 40: 243-267.
119. Mohan S, Hsiao C, Bowman JC, Wartell R, Williams LD (2010) RNA tetraloop folding reveals tension between backbone restraints and molecular interactions. J Am Chem Soc 132: 12679-12689.
120. Brion P, Westhof E (1997) Hierarchy and dynamics of RNA folding. Annu Rev Biophys Biomol Struct 26: 113-137.
121. Mortimer SA, Weeks KM (2008) Time-Resolved RNA SHAPE Chemistry. J Am Chem Soc 130: 16178-16180.
122. Athavale SS, Petrov AS, Hsiao C, Watkins D, Prickett CD, et al. (2012) RNA Folding and Catalysis Mediated by Iron (II). PLoS ONE 7: e38024.
123. Bernstein DS, Buter N, Stumpf C, Wickens M (2002) Analyzing mRNA-protein complexes using a yeast three-hybrid system. Methods 26: 123-141.
124. Hook B, Bernstein D, Zhang B, Wickens M (2005) RNA-protein interactions in the yeast three-hybrid system: affinity, sensitivity, and enhanced library screening. RNA 11: 227-233.
125. Stumpf CR, Opperman L, Wickens M (2008) Analysis of RNA-protein interactions using a yeast three-hybrid system. Methods Enzymol 449: 295-315.
126. Job P (1928) Studies on the formation of complex minerals in solution and on their stability. Annales De Chimie France 9: 113-203.
127. Cantor C, Schimmel P (1984) Biophysical Chemistry (I-III). New York: Academic Press.
128. Prodromou C, Pearl LH (1992) Recursive PCR - a Novel Technique for Total Gene Synthesis. Protein Engineering 5: 827-829.
129. Sandhu GS, Aleff RA, Kline BC (1992) Dual Asymmetric PCR - One-Step Construction of Synthetic Genes. BioTechniques 12: 14-16.
130. Sampson JR, Uhlenbeck OC (1988) Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. Proc Natl Acad Sci U S A 85: 1033-1037.
131. Wurster SE, Bida JP, Her YF, Maher LJ, 3rd (2009) Characterization of anti-NF-kappaB RNA aptamer-binding specificity in vitro and in the yeast three-hybrid system. Nucleic Acids Res 37: 6214-6224.

132. Cassiday LA, Maher LJ, 3rd (2001) In vivo recognition of an RNA aptamer by its transcription factor target. Biochemistry 40: 2433-2438.
133. Barkholt V, Jensen AL (1989) Amino acid analysis: determination of cysteine plus half-cystine in proteins after hydrochloric acid hydrolysis with a disulfide compound as additive. Anal Biochem 177: 318-322.
134. Recht MI, Fourmy D, Blanchard SC, Dahlquist KD, Puglisi JD (1996) RNA sequence determinants for aminoglycoside binding to an A-site rRNA model oligonucleotide. Journal of Molecular Biology 262: 421-436.
135. Grilley D, Soto AM, Draper DE (2006) Mg2+–RNA interaction free energies and their relationship to the folding of RNA tertiary structures. Proceedings of the National Academy of Sciences 103: 14003-14008.
136. Zengel JM, Jerauld A, Walker A, Wahl MC, Lindahl L (2003) The extended loops of ribosomal proteins L4 and L22 are not required for ribosome assembly or L4-mediated autogenous control. RNA 9: 1188-1197.
137. Diner EJ, Hayes CS (2009) Recombineering reveals a diverse collection of ribosomal proteins L4 and L22 that confer resistance to macrolide antibiotics. Journal of Molecular Biology 386: 300-315.
138. Evguenieva-Hackenberg E (2005) Bacterial ribosomal RNA in pieces. Molecular microbiology 57: 318-325.
139. Kramer G, Rauch T, Rist W, Vorderwülbecke S, Patzelt H, et al. (2002) L23 protein functions as a chaperone docking site on the ribosome. Nature 419: 171-174.
140. Akanuma G, Nanamiya H, Natori Y, Yano K, Suzuki S, et al. (2012) Inactivation of ribosomal protein genes in Bacillus subtilis: importance of each ribosomal protein for cell proliferation, as well as cell differentiation. J Bacteriol.
141. Huber D, Rajagopalan N, Preissler S, Rocco MA, Merz F, et al. (2011) SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria. Mol Cell 41: 343-353.
142. Wilson DN, Beckmann R (2011) The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. Current opinion in structural biology 21: 274-282.
143. Schulze H, Nierhaus KH (1982) Minimal set of ribosomal components for reconstitution of the peptidyltransferase activity. EMBO J 1: 609-613.
144. Agmon I (2009) The dimeric proto-ribosome: Structural details and possible implications on the origin of life. Int J Mol Sci 10: 2921-2934.
145. Smith TF, Lee JC, Gutell RR, Hartman H (2008) The origin and evolution of the ribosome. Biol Direct 3: 16.
146. Sloof P, Van den Burg J, Voogd A, Benne R, Agostinelli M, et al. (1985) Further characterization of the extremely small mitocbondrial ribosomal RNAs from trypanosomes: a detailed comparison of the 9S and 12S RNAs from Crithidia fasciculate and Trypanosoma brucei with rRNAs from other organisms. Nucleic Acids Res 13: 4171-4190.